Comments of General Electric Company on the Ecological Risk Assessment for the General Electric/ Housatonic River Site, Rest of River (July 2003 Draft)

Prepared by:

BBL Sciences
Petaluma, California

ARCADIS G&M, Inc.
Portland, Maine

Branton Environmental Consulting
Vancouver, British Columbia

LWB Environmental Services
Oak Ridge, Tennessee

On Behalf of:

General Electric Company

September 29, 2003

TABLE OF CONTENTS

ATTACHMENTS
ACRONYMS
EXECUTIVE SUMMARY i
1. INTRODUCTION 1-1
2. CHARACTERIZATION OF ECOSYSTEM (Question 1) 2-1
2.1 Overview of Site 2-2
2.2 Characterization of Ecosystem in ERA 2-4
3. SELECTION OF ASSESSMENT AND MEASUREMENT ENDPOINTS (Question 2) 3-1
3.1 Selection of Assessment Endpoints 3-2
3.2 Selection of Measurement Endpoints 3-4
4. GENERAL ISSUES RELATING TO EVALUATION OF ASSESSMENT ENDPOINTS (Question 3) 4-1
4.1 Evaluation of Field Studies 4-2
4.2 Evaluation of HQs Based on Modeled Exposures and Effects 4-4
4.3 Other Issues in Weight-of-Evidence Evaluations 4-6
4.4 Incomplete Documentation of Key Studies 4-7
5. BENTHIC INVERTEBRATES (Question 3.1) 5-1
5.1 EPA Benthic Community Field Study 5-2
5.2 Site-Specific Toxicity Studies 5-4
5.3 Comparison to Literature-Based Benchmarks or Effects Levels 5-8
5.3.1 Comparisons to generic sediment and water quality values 5-8
5.3.2 Derivation of tissue-based threshold concentration 5-8
5.4 Overall Assessment 5-9
5.5 Extrapolation to Downstream Reaches 5-10
6. AMPHIBIANS (Question 3.2) 6-1
6.1 Site-Specific Studies on Leopard Frogs 6-2
6.1.1 EPA's leopard frog study 6-2
6.1.2 GE leopard frog egg mass survey 6-5
6.1.3 EPA's anecdotal observations 6-6
6.2 Site-Specific Studies on Wood Frogs 6-7
6.2.1 EPA's wood frog study 6-7
6.2.2 GE's wood frog study 6-10
6.2.3 EPA's population model 6-10
6.3 EPA's Vernal Pool Study 6-11
6.4 Unrealistically Conservative Effects Thresholds 6-12
6.5 Overall Assessment 6-13
6.6 Extrapolation to Downstream Reaches 6-14
7. FISH (Question 3.3) 7-1
7.1 EPA's Site-Specific Toxicity Studies 7-2
7.1.1 Lack of consistent evidence of dose-dependent PCB effects 7-3
7.1.2 No dose-response relationship between PCBs and swim bladder deformities 7-4
7.2 Site-Specific Fish Community and Population Studies 7-5
7.3 Tissue-Based Effects Thresholds 7-7
7.3.1 Effects metrics based on site-specific toxicity studies 7-8
7.3.2 Effects thresholds derived from literature 7-8
7.4 Overall Assessment 7-9
7.5 Extrapolation of Risks to Downstream Reaches 7-10
8. INSECTIVOROUS BIRDS (Question 3.4) 8-1
8.1 EPA's Site-Specific Field Study of Tree Swallows 8-2
8.2 GE's Site-Specific Field Study of Robins 8-3

8.3 Modeled Exposures and Effects 8-4
8.3.1 Modeled exposures and effects for tree swallows 8-5
8.3.2 Modeled exposures and effects for American robins 8-6
8.4 Overall Assessment 8-7
9. PISCIVOROUS BIRDS (Question 3.5) 9-1
9.1 GE's Site-Specific Field Study on Kingfishers 9-2
9.2 Selection of Ospreys as Representative Receptor 9-4
9.3 Modeled Exposures and Effects 9-7
9.3.1 Exposure assumptions 9-7
9.3.2 Effects metrics 9-8
9.3.3 Correction of Hazard Quotients 9-9
9.4 Overall Assessment 9-9
10. PISCIVOROUS MAMMALS (Question 3.7) 10-1
10.1 Mink Feeding Study 10-2
10.2 Jaw Lesion Study 10-5
10.3 EPA and GE Mink and River Otter Surveys 10-6
10.3.1 EPA mink and otter survey 10-6
10.3.2 GE mink and otter survey 10-7
10.4 Modeled Exposure and Effects 10-11
10.5 Overall Assessment 10-13
10.6 Extrapolation to Downstream Reaches 10-14
11. OMNIVOROUS AND CARNIVOROUS MAMMALS (Question 3.6) 11-1
11.1 EPA Field Surveys 11-2
11.2 GE Population Demography Field Study of Short-tailed Shrews 11-3
11.3 Modeled Exposure and Effects 11-5
11.4 Overall Assessment 11-6
12. THREATENED AND ENDANGERED SPECIES (Question 3.8) 12-1
12.1 Modeled Exposures and Effects for Bald Eagles 12-2
12.2 Modeled Exposures and Effects for American Bitterns 12-4
12.3 Modeled Exposures and Effects for Small-Footed Myotis 12-5
12.4 Overall Assessment 12-5
12.5 Extrapolations of Bald Eagle Risks to Downstream Reaches 12-8
13. DISCUSSIONS AND CONCLUSIONS (Question 4) 13-1
13.1 Inconsistent Weight-of-Evidence Evaluations 13-2
13.2 Overemphasis on HQs 13-5
13.3 Extrapolations to Downstream Reaches 13-6
13.4 Extrapolations to Other Species 13-7
13.5 Ecological Implications 13-7
13.6 Overall Uncertainty Analysis 13-8
13.7 Overall Conclusions 13-9
14. REFERENCES 14-1

LIST OF TABLES

ES-1	Summary of Conclusions for Receptor Groups	xvi
5-1	Recalculation of Effects Thresholds Using Synoptic Data and Excluding Redundant Endpoints	5-7
5-2	Assessment Endpoint: Community Structure, Survival, Growth, and Reproduction of Benthic Invertebrates	5-11
6-1	Assessment Endpoint: Community Condition, Survival, Reproduction, Development, and Maturation of Amphibians	6-19
7-1	Assessment Endpoint: Survival, Growth, and Reproduction of Fish	7-11
8-1	Assessment Endpoint: Survival, Growth, and Reproduction of Insectivorous Birds	8-9
9-1	Assessment Endpoint: Survival, Growth, and Reproduction of Piscivorous Birds	9-12
10-1	Assessment Endpoint: Survival, Growth, and Reproduction of Piscivorous Mammals	10-15
11-1	Relationship Between the Survival of *Blarina brevicauda* on Six Live-Trapping Grids from Summer to Fall 2001 and Spatially-Weighted Arithmetic tPCB Concentrations (as Calculated in the ERA)	11-5
11-2	Assessment Endpoint: Survival, Growth, and Reproduction of Carnivorous and Omnivorous Mammals	11-8
12-1	Comparison of ERA's Weight-of-Evidence Evaluations for Avian HQ for tPCBs	12-7
12-2	Assessment Endpoint: Survival, Growth, and Reproduction of Threatened and Endangered Species	12-9

LIST OF FIGURES

6-1	Regression of Net Abnormality-Free Metamorph Output vs. Tissue PCB Concentrations	6-9
6-2a,b,c	Wetlands and Vernal Pools within the 100-Year Floodplain of the Housatonic River, Downstream of Woods Pond	6-16
7-1	Fish Community Structure in the Mainstem of the Housatonic River	7-6
8-1	Numbers of Robins Fledged per Nest	8-3
8-2	PCB Exposure in Target and Reference Robins	8-4
8-3	Impact of Alternative Assumptions on tPCB Hazard Quotients for Insectivorous Birds	8-5
9-1	Survival of Belted Kingfisher Nestlings	9-3
9-2	Great Blue Heron Productivity in Massachusetts (1980-1990)	9-6
9-3	Impact of Alternative Assumptions on tPCB Hazard Quotients for Piscivorous Birds	9-9
10-1	Mink Kit Survival at Six Weeks	10-4
10-2	Mink Photographed on the Mainstem Using Motion Detector Camera in Winter 2003	10-8
10-3	River Otter Photographed on the Mainstem Using Motion Detector Camera in Winter 2003	10-8
10-4	Mink Track Locations (Winter 2003)	10-9
10-5	Impact of Alternative Assumptions of tPCB Hazard Quotients for Piscivorous Mammals	10-13
11-1	Population Densities (+ 1 SE) per ha of *Blarina brevicauda* in 2001 on Six Live-Trapping Grids Along the Housatonic River, Massachusetts between Pittsfield and Woods Pond	11-3
12-1	Impact of Alternative Assumptions on tPCB Hazard Quotients for Threatened and Endangered Birds	12-4

ATTACHMENTS (Separately Bound)

A. Errors and Inconsistencies in the ERA

B. Critique of the Method Used to Calculate the 95% Upper Confidence Limit on the Mean of the Sampling Data

C. Evaluation of the Benthic Community Study Presented in the ERA

D. Evaluation of Effects Metrics Developed from Benthic Invertebrate Bioassay Data

E. Evaluation of EPA's Northern Leopard Frog Reproduction and Development Study

F. Northern Leopard Frog (*Rana pipiens*) Egg Mass Survey

G. Evaluation of EPA's Wood Frog Vernal Pool Study and EPA's Wood Frog Population Model

H. Experimental Analysis of the Context-Dependent Effects of Early Life-Stage PCB Exposure on *Rana Sylvatica*

I. Evaluation of USGS Phase I and Phase II Fish Toxicity Studies and Effects Metrics for Fish

J. *In Situ* Reproduction, Abundance and Growth of Young-of-Year and Adult Largemouth Bass in a Population Exposed to Polychlorinated Biphenyls

K. *Productivity of American Robins Exposed to Polychlorinated Biphenyls, Housatonic River, Massachusetts, USA* (Manuscript In Press)

L. Productivity and Density of Belted Kingfishers on the Housatonic River

M. Productivity Data for Great Blue Herons Breeding in Massachusetts (1980-1999)

N. Analysis of Kit Survivability Data from EPA's Mink Feeding Study

O. Evaluation of Piscivorous Mammals – Presence/Absence, Distribution, and Abundance in the Housatonic River Floodplain

P. Comments Relating to Mink and Otter Field Studies

Q. *Demography of Short-Tailed Shrew Populations Living on Polychlorinated Biphenyl-Contaminated Sites*

R. Comments Relating to Short-Tailed Shrew Demography Study

ACRONYMS

ºC	degrees Celsius
95% UCL	95% upper confidence limit on the mean
Ah-R	aryl hydrocarbon receptor
CBS	Carolina Biological Supply
COC	chemical of concern
COPC	chemical of potential concern
CPUE	catch per unit effort
ECOD	ethoxycoumarin-O-deethylase
ED20/ED50	concentration at which the frequency of observed effects is 20% or 50%
EPC	exposure point concentration
EPA	U.S. Environmental Protection Agency
ERA	*Ecological Risk Assessment for General Electric (GE)/Housatonic River Site, Rest of River*
EROD	ethoxyresorufin-O-deethylase
FEL	Fort Environmental Laboratory
g	gram
GE	General Electric Company
GPS	global positioning system
ha	hectare
HQ	hazard quotient
HSI	Habitat Suitability Index
LOAEL	lowest observed adverse effect level
MassGIS	Massachusetts Geographic Information System
MATC	maximum acceptable threshold concentration
MDFW	Massachusetts Division of Fish and Wildlife
mg/kg	milligrams per kilogram (equivalent to parts per million)

mg/kg bw/d	milligrams per kilogram body weight per day
MSD	minimum significant difference
ng/g	nanograms per gram (equivalent to parts per billion)
ng/kg	nanograms per kilogram (equivalent to parts per trillion)
ng/kg bw/d	nanograms per kilogram body weight per day
NOAEL	no observed adverse effect level
PAH	polycyclic aromatic hydrocarbon
PCB	polychlorinated biphenyl
PSA	Primary Study Area
SQV	Sediment Quality Value
SVOC	semi-volatile organic compound
TCDD	tetrachlorodibenzo-p-dioxin
T&E	threatened and endangered
TEQ	dioxin toxicity equivalent
TIE	Toxicity Identity Evaluation
TOC	total organic carbon
tPCBs	total polychlorinated biphenyls
USGS	U.S. Geologic Survey
WWTP	wastewater treatment plant
YOY	young-of-year

EXECUTIVE SUMMARY

The General Electric Company (GE) is providing these Comments to the U.S. Environmental Protection Agency (EPA) and the Peer Review Panel on EPA's public review draft of the *Ecological Risk Assessment for General Electric (GE)/Housatonic River Site, Rest of River* (ERA). The objective of these Comments is to present additional information, viewpoints, and analyses relating to the ERA's characterization of the potential risks to ecological receptors due to polychlorinated biphenyls (PCBs) and dioxin toxicity equivalents (TEQs) in the Rest of River portion of the Housatonic River and its floodplain. These Comments generally follow the structure of the Peer Review Charge for the ERA. They discuss in the text the main points that GE and its consultants have identified relating to the ERA's description, interpretation, evaluation, and weighting of the various lines of evidence and the conclusions drawn in the ERA. They focus, in particular, on instances in which we believe that the ERA is not objective, consistent, or reasonable (as specified in the Peer Review Charge) and is not accurate, reliable, and unbiased (as required by EPA [2002a] guidelines). These points are supported by a series of attachments that contain more detailed evaluations of some of EPA's studies and more detailed descriptions of the studies sponsored by GE. (The following headings in this Executive Summary reference in parentheses the corresponding Peer Review Charge questions, as well as the sections of the Comments where the points are discussed in greater detail.)

Characterization of Ecosystem (Charge Question 1; Comments Section 2)

The ecological characterization that was conducted for the ERA provides substantial information on the ecosystem and the fish and wildlife populations present at the site, particularly in the Primary Study Area (PSA), which includes the Housatonic River and its floodplain from the confluence of the East and West Branches of the river in Pittsfield to Woods Pond Dam in Lenox/Lee. PCBs have been present in this system at elevated concentrations, especially in the PSA, since the 1930s, and hence local fish and wildlife populations have been exposed to elevated PCB concentrations for many generations. Nevertheless, the ecological surveys conducted for the ecological characterization found no obvious population-level impacts. Rather, they showed abundant, diverse, and thriving fish and wildlife populations and communities in the Housatonic River and its floodplain, including numerous species of invertebrates, fish, reptiles, amphibians, birds, and mammals.

Based on this ecological characterization, the ERA generally provides an adequate description of the local ecosystem and appropriately applies that information in the problem formulation, with a few specific exceptions. Those exceptions, discussed further below, relate to the ERA's inadequate discussion of habitat limitations for amphibians and bald eagles in areas downstream of the PSA and to its selection of ospreys, which do not breed in the area, as a representative species.

Selection of Assessment and Measurement Endpoints (Charge Question 2; Comments Section 3)

EPA (1999) guidance specifies that assessment endpoints in an ERA should address local populations and communities of biota, rather than individual organisms (with the possible exception of threatened and endangered [T&E] species). While effects on such populations and communities can be extrapolated from studies on individuals or groups of individuals, the assessment endpoints should be focused on the local populations and communities. Two of the eight assessment endpoints selected in the ERA (for benthic invertebrates and amphibians) do include one community-level endpoint each; however, the remaining specific endpoints for those two receptors and the six other assessment endpoints all focus on survival, growth, and reproduction of individual organisms, rather than directly on population- or community-level parameters. To be consistent with EPA guidance, these assessment endpoints should not be evaluated in isolation – i.e., on an individual-level basis. Rather, they are appropriate *only* if the survival, growth, and reproduction effects are evaluated *in terms of their impacts on local populations and communities*. For the most part, however, the ERA focuses and bases its conclusion on individual-level effects, without adequately considering their potential impact on the local populations and communities. For example, for most receptor groups, the ERA does not attempt to assess the implications of the individual-level effects for local populations or communities. In addition, the ERA generally assigns greater weight to the studies of individual-level effects than to the available site-specific data that directly address local populations or communities. These approaches are inconsistent with the proper focus of an ecological risk assessment and with EPA (1999) guidance.

With respect to the measurement endpoints selected in the ERA to evaluate the assessment endpoints, GE believes that most of those measurement endpoints are, in concept, appropriate types of studies and analyses, provided that their results are evaluated in terms of their relevance to local populations and communities. In many cases, however, the ERA does not provide such an evaluation. Further, in many cases, the ERA's interpretation and/or weighting of the specific studies and analyses, as well as the conclusions drawn from them, are unbalanced, inconsistent, or not well supported by the evidence, as will be discussed below.

Evaluation of Assessment Endpoints – General Issues (Charge Question 3; Comments Section 4)

There are a number of general issues that affect the ERA's evaluation of multiple assessment endpoints and its overall approach to the evaluation of risk. These relate mainly to the ERA's application of the weight-of-evidence approach, its general evaluation of field studies, and its general evaluation of Hazard Quotients (HQs) based on modeled exposures and effects.

It is reasonable to apply a weight-of-evidence approach to resolve multiple lines of evidence. However, when this approach is applied qualitatively, as it is in the ERA, there are many opportunities for subjective judgments. In this case, where EPA has been the judge of its own studies and those sponsored by GE, there is an inherent potential for bias. Further, in the interpretation or weighting of the various lines of evidence, there were many opportunities for interpreting the data or weighting the lines of evidence in a manner that favors a conclusion of risk. Thus, it is critical for the peer reviewers to closely examine the weights, evidence of harm, and magnitude of response reported for each measurement endpoint in the ERA, as well as the conclusions reached, to evaluate whether they are objective, consistent, and reasonable (as specified in the Charge) and are accurate, reliable, and unbiased (as required by EPA [2002a] guidance).

Numerous field studies were conducted in connection with the ERA, both by EPA contractors and experts and by contractors and experts retained by GE. Such field studies have many inherent advantages over laboratory tests and model-based predictions, because they can provide site-specific and species-specific data on the actual condition of a local population of a given type of receptor that has been exposed to PCBs over many generations. Further, field studies can examine the responses of such populations to PCBs against the backdrop of natural population dynamics, reflecting the influences of density-dependence, predation, competition, and fluctuating climatic conditions. Nevertheless, the ERA assigns *unduly low weight to many of the field studies*, particularly relative to the weight given to modeled exposures and effects based on the literature.

By contrast, the ERA places too much weight on HQs based on modeled exposures and effects. These HQs are *overly conservative* because they:

- Do not address population-level impacts, but rather focus on individual-level impacts;
- Use inappropriate exposure assumptions for wildlife receptors, including: (a) modeled generic food intake rates when measured species-specific rates are available in the literature for most species evaluated; and (b) in some cases, unrealistic assumptions for foraging time within the PSA;
- Are mostly based on effects metrics derived from laboratory studies on other species and/or using PCB mixtures different from those in the PSA; and
- Use overestimated exposure point concentrations based on either the maximum detected concentration or the 95% upper confidence limit (UCL) on the mean, calculated using a statistical procedure (Land's H-statistic) that overstates the upper bound on the mean.

The over-conservatism of these HQs is evident from the fact that, for receptors for which both HQs and site-specific field data are available, the HQs consistently indicate substantially more risk than the site-specific data, often predicting intermediate or high risk where the field studies showed no evidence of harm.

Finally, in addition to the above issues, the ERA has *misapplied the weight of evidence approach* by: (1) incorrectly characterizing evidence of harm as "undetermined" for several studies that in fact showed no evidence of harm; and (2) inappropriately assigning a magnitude of response to certain endpoints with "undetermined" evidence of harm, which is illogical.

Benthic Invertebrates (Charge Question 3.1; Comments Section 5)

The ERA concludes that there is a significant risk of adverse effects to benthic invertebrates due to PCBs throughout the PSA and in some downstream areas, with a moderate to high level of certainty. This conclusion is based on a misinterpretation or overstatement of the underlying lines of evidence.

<u>EPA's site-specific benthic invertebrate community study.</u> This study found no adverse effects of PCBs on the benthic invertebrate community at sites with fine-grained sediments. While it did report significant differences in benthic invertebrate parameters between PSA sites with coarse-grained sediments and coarse-grained reference sites, that observation does not show that PCBs are having adverse effects at those sites, because: (a) the study did not adequately take into account the differences in habitat factors among the sites (e.g., presence of sand) that may account for the observed differences in community metrics; (b) there was no correlation between PCB concentrations in the sediments and those in the invertebrates' tissue; and (c) the study showed no significant exposure-response relationship between PCB concentrations in the sediments (over a wide range) and any of the benthic community metrics. Moreover, GE's independent multiple regression analysis of the data shows that sediment PCB concentrations accounted for only a very small fraction of the variability in community metrics at the coarse-grained sites and thus have *no meaningful influence on the benthic community structure at those sites*.

<u>EPA's laboratory and in situ toxicity tests using site-specific sediments.</u> The ERA uses the data from these toxicity tests to develop a Maximum Acceptable Threshold Concentration (MATC) for PCBs of 3 mg/kg in sediments, which is then used to assess risks within and downstream of the PSA. However, the test results do not support that sediment MATC because the method used to calculate it is flawed in several respects. Specifically, that method: (a) used PCB concentration data collected from sediments over a several-month period, rather than from the sediments to which the invertebrates were actually

exposed; (b) included comparisons to a laboratory control despite differences between that control and reference sediments that cannot be explained on the basis of PCB concentrations; (c) used an averaging method that double counts certain endpoints; and (d) included 20% effects values that likely overestimate effects. A reanalysis of the toxicity test data, correcting for these flaws, would result in a *sediment PCB MATC of approximately 8 mg/kg*. Even that threshold, however, would *likely overstate risks* due to the use of questionable reference sites and the possible confounding influence of other factors (e.g., other chemicals). This is demonstrated by the fact that the same stations from which sediments were collected for these toxicity tests (which were fine-grained sites) showed no evidence of effects in the benthic community study despite sediment PCB levels considerably higher than 8 mg/kg (for some individual samples up to 50 mg/kg).

<u>Comparisons to benchmarks or effects levels derived from the literature.</u> The ERA also compares sediment and surface water concentrations to generic Sediment Quality Values and water quality criteria. Such comparisons are inappropriate since those values do not reflect site-specific conditions and since site-specific data are available. In addition, the ERA derives a tissue-based threshold PCB concentration of 3 mg/kg based on the literature and uses that value for comparison as well. That threshold concentration, however, is unsupported because it is not consistent with the site-specific data and does not recognize the limitations of the underlying literature studies.

<u>Overall assessment.</u> *The ERA overstates the magnitude of risks to benthic invertebrates in the PSA and the certainty of the conclusions. Although EPA's toxicity test data, properly interpreted, would support a sediment MATC of around 8 mg/kg PCBs, that threshold likely overestimates risks, as shown by the fact that there were no adverse PCB-related effects on benthic invertebrate community parameters in the community field study, even at sites with considerably higher PCB concentrations.*

Amphibians (Charge Question 3.2; Comments Section 6)

The ERA concludes that amphibians have a high risk of ecologically significant effects due to PCBs throughout the PSA and in downstream areas, with a moderate to high level of certainty. Again, this conclusion is based on misinterpretations or over-interpretations of several of the lines of evidence.

<u>Site-specific leopard frog studies.</u> The ERA states that EPA's 2001 leopard frog toxicity study showed impaired reproductive condition in adult frogs from the PSA (e.g., high rates of immature eggs in females), as well as developmental effects in exposed larvae. That conclusion, however, cannot be supported because it relies to a large extent on comparisons to frogs purchased from a commercial distributor, which were collected in very different locations, may have been in different reproductive

stages, and were subjected to less handling stress (including much shorter holding times) than frogs from the PSA. Indeed, the study authors recognized this and (unlike the ERA) excluded the comparisons to those commercially purchased frogs from their statistical analyses. Moreover, the adult females from the PSA may have been collected at the wrong time to capture a representative sample of reproductively mature females.

Given the finding of the EPA study regarding the impaired reproductive condition of adult leopard frogs in the PSA, a survey was conducted in 2003 on behalf of GE to assess the presence and extent of leopard frog egg masses in the PSA. This survey found 216 egg masses in 17 PSA ponds, thus showing *significant reproductive activity in the leopard frog population in the PSA* and suggesting that the EPA study finding was likely due to some factor other than PCBs (perhaps poor timing in the collections).

<u>Site-specific wood frog studies.</u> The ERA also puts heavy reliance on EPA's 2001 three-phase wood frog toxicity study. Although that study collected data on a wide range of endpoints relevant to the survival, development, and maturation of wood frog egg masses, larvae, and metamorphs, the ERA relies only on the few endpoints that showed a significant PCB effect – larval malformations and metamorph abnormalities in Phases I and II and skewed sex ratio (more females than males) in Phase III. However, the ERA fails to consider that, in the same study, PCB exposure was not related to endpoints that are more directly relevant to the local populations, such as survival and metamorphosis. In fact, GE's independent analysis shows that there was no significant relationship between PCB exposure and net (abnormality-free) metamorph output in these data, thus indicating that *the rates of abnormalities observed did not affect overall reproductive output* and thus would not result in reduced recruitment of metamorphs into the population. Moreover, field data collected by EPA indicate that the sex ratios of breeding adult wood frogs were not skewed. Further, even in metamorphs, analysis of the data shows no significant relationship between tissue PCB concentrations (which provide a more direct measure of dose than sediment PCB concentrations) and sex ratio; and in any event, sex ratio in amphibians can be affected by a number of environmental factors other than PCBs.

The ERA also relies on EPA's vernal pool study, which involved four PSA ponds and reported low species density and richness in ponds with higher PCB concentrations and high malformation rates in larval wood frogs in all the ponds. That study does not provide reliable evidence of PCB effects because it involved only four ponds and failed to consider potential confounding factors (such as pond size). In addition, this study showed no exposure-response relationship for malformations across ponds with a wide range of sediment PCB concentrations (0.7 to 32 mg/kg), and showed low rates of malformations in metamorphs and adult wood frogs, indicating that larval malformations do not translate to later developmental stages.

EPA's wood frog population model. To assess potential population-level impacts, EPA developed a population model for wood frogs. That model is flawed, primarily due to its failure to consider density-dependent processes. As a result, it predicts a median time to extinction of 32 years even in the absence of PCB exposure and 17-30 years with PCB exposure – results that are refuted by the presence of wood frog populations in the PSA despite the decades of PCB presence there.

Effects thresholds. The ERA develops a sediment/soil MATC of 3 mg/kg PCBs based primarily on the sex ratio data from EPA's wood frog study and a tissue MATC of 1 mg/kg PCBs based on those data as well as on review of the literature. As noted above, however, the sex ratio data from EPA's wood frog study do not provide a reasonable basis for establishing an effects threshold. Further, the literature-based threshold is based on a misinterpretation of the underlying studies. The malformation data from EPA's wood frog study would support PCB MATCs of 38 mg/kg in sediment/soil and 6 mg/kg in frog tissue. Even those thresholds, however, would likely overstate population risks because there is no evidence that the malformation rates affect survival, growth, or net reproductive output of healthy (abnormality-free) metamorphs.

Overall assessment. *Overall, while there is some evidence of increased larval and metamorph abnormalities among amphibians in the PSA, the available data do not indicate that these abnormalities translate into adverse effects on local amphibian populations.*

Extrapolation to reaches downstream of the PSA. The ERA also overestimates risk to amphibians in downstream reaches by relying on an overly conservative sediment/soil MATC (as noted above) and also by applying this MATC to the entire 100-year floodplain below the PSA, when publicly available mapping indicates that only a small portion of that floodplain contains wetlands or vernal pools that provide suitable habitat for amphibian larvae (the life stage on which the MATC is based).

Fish (Charge Question 3.3; Comments Section 7)

The ERA concludes, with moderate confidence, that there are ecologically significant but low-magnitude risks to fish in the PSA from both PCBs and TEQs, and that there are marginal and uncertain risks to coldwater fish (but not warmwater fish) downstream of the PSA. Based on review of the underlying lines of evidence, these conclusions overstate the risks to fish in the Housatonic River.

EPA's site-specific toxicity studies. The ERA concludes that EPA's Phase I and Phase II toxicity studies provide evidence of PCB-related toxicosis in fish. However, a careful review of the data from those studies shows that they *did not show consistent relationships between PCB exposure and adverse effects*. While these studies did find various statistically significant relationships between at least one Housatonic

River site and the reference site for a number of adult and offspring endpoints, those differences were not consistent among the Housatonic River sites or among developmental phases or between the Phase I and II studies, and did not show clear exposure-response relationships with PCBs. In fact, to calculate effects thresholds, the ERA inappropriately combines all mortality and abnormality data from Phase II, when most of those individual endpoints did not show an exposure-response relationship with PCBs.

Site-specific fish population and community studies. Field studies conducted by both EPA and GE showed that there are *self-sustaining populations and communities of fish in the Housatonic River, including the PSA.* Further, GE's study demonstrated that the largemouth bass population in the PSA is reproducing and has population parameters within the range that would be expected for a similar system in the Northeast. It also showed that the tributaries to the PSA portion of the river have very limited or unsuitable habitat for largemouth bass. Hence, the only way that the PSA portion of the river could support such a healthy largemouth bass population is through internal reproduction. In fact, this study documented the presence of numerous (> 70) active largemouth bass nests in the PSA. Accordingly, any individual-level PCB effects that may be occurring in largemouth bass in the PSA do not appear to be affecting the sustainability of the bass population. In short, given that PCBs have been present in the PSA for many decades, these studies show that the PCBs are not adversely affecting the overall populations and communities of fish in the river. In these circumstances, the ERA incorrectly asserts that these studies showed "undetermined" evidence of harm. In fact, they found no evidence of harm.

Tissue effects thresholds. The ERA derives tissue-based effects thresholds for fish based both on its two-phase toxicity studies and on its review of the literature. These thresholds are overly conservative. For example, the thresholds derived from the toxicity studies are overstated for the same reasons discussed above (e.g., lack of consistent exposure-response relationships with PCBs for individual endpoints) and because they rely on unsupported conversion factors (for converting from egg to whole body tissue concentrations and for extrapolating from warmwater to coldwater species). The literature-based effects thresholds are overstated because they: (a) inappropriately include studies of a highly sensitive species (lake trout) that is not found in the Housatonic River; (b) are based (for tPCBs) on an inappropriate averaging of different effects metrics in different tissue types; and (c) utilize an egg to whole body conversion factor which, based on the overall literature, is unnecessary.

Overall assessment. *The available data do not indicate that there are ecologically significant risks to fish from PCBs and TEQs in the Housatonic River in terms of population sustainability. Rather, they indicate that, while some individual-level effects may be occurring, there are negligible risks to the overall fish populations and communities.*

Insectivorous Birds (Charge Question 3.4; Comments Section 8)

The ERA assesses risks to insectivorous birds based on field studies and model-based HQ analyses of tree swallows and American robins. It concludes that risks to insectivorous birds from PCBs and TEQs in the PSA are low, but that this conclusion is uncertain due to conflicting outcomes for the field studies (showing no significant effects) and HQs (predicting intermediate to high risks).

Site-specific field studies. As the ERA recognizes, both EPA's field study of tree swallows and GE's field study of robins showed *no evidence that exposure to PCBs in the PSA had an adverse effect on the reproductive success of these birds* nesting in the PSA. These were high-quality studies, and both should be given high weight. (In fact, based on EPA's tree swallow study, a site-specific dose-based effects metric can be calculated and should be used in the HQ analyses for *other* avian species for which species-specific effects data are not available, as the effects metric for the most tolerant avian species.)

Modeled exposures and effects. The HQ analyses unnecessarily rely on modeled and literature-based inputs when site-specific data are available. For example, the HQs for tree swallows are based on modeled tissue PCB concentrations for nestlings, when the site-specific tissue-based effects metrics for PCBs are based on pippers (a different age class) and measured site-specific tissue PCB data on pippers are available. Similarly, the HQs for robins use literature-based effects metrics for other species, rather than effects metrics derived from the site-specific studies. As a result, the HQs *substantially overestimate risk*, as evidenced by the fact that the risks they predict are contradicted by the field data. Correcting the HQs to use site-specific inputs would yield substantially lower HQ results (less than 1 for both species) that correspond more closely with the results of the field studies.

Overall assessment. *A proper interpretation of the data, with the correction of the HQ analyses, indicates, with high certainty, that insectivorous birds are at negligible risk from exposure to PCBs and TEQs in the PSA.*

Piscivorous Birds (Charge Question 3.5; Comments Section 9)

The ERA assesses risks to piscivorous birds based on a field study of belted kingfishers and model-based HQs for belted kingfishers and ospreys. It concludes that risks to kingfishers from PCBs in the PSA are low, with uncertainty due to inconsistent results between the field study and the HQ analyses, and that risks to ospreys from PCBs in the PSA are intermediate to high, with uncertainty due to the availability of only one line of evidence. GE's principal concerns with these conclusions relate to the inappropriate selection of ospreys as a representative species and the use of overly conservative assumptions in the HQ analyses for both species.

Site-specific field study on kingfishers. This GE study found *no evidence of harm to belted kingfishers* – the most highly exposed piscivorous bird known to breed in the PSA. It showed that the local population is breeding successfully and that the density of the population is consistent with the quality of the available habitat.

Selection of ospreys and use of great blue herons as alternative. Breeding ospreys should not be included in the ERA as a representative receptor because all available data indicate that *ospreys do not breed in Western Massachusetts.* Since kingfishers have a higher exposure potential than any other piscivorous bird known to breed in the PSA, they suffice as a representative species for this feeding guild. If, however, a second representative species is needed, then *great blue herons would be a more appropriate choice* than ospreys, since great blue herons do breed within foraging distance of the Housatonic River. Field data are available showing no difference in productivity between great blue herons breeding within foraging distance of the river and those breeding elsewhere in Massachusetts. These data could serve as a second line of evidence for this species, in addition to an appropriate HQ analysis.

Modeled exposures and effects. The model-based HQs substantially overestimate risks to both belted kingfishers and ospreys for two reasons. First, they use overly conservative exposure assumptions, including: (a) modeled food intake rates when measured rates are available in the literature; and (b) for ospreys, the unrealistic assumption that ospreys would obtain 100% of their prey from the PSA, when any ospreys foraging in the area would likely be present for only a few days during migration. Second, they use inappropriate effects metrics (based on chickens and kestrels for PCBs) when more supportable and relevant effects metrics are available. The excess conservatism of the HQ analyses is illustrated by their prediction of high risks to kingfishers when the field study found no evidence of harm.

Overall assessment. *A proper evaluation of the evidence, including use of kingfishers alone or kingfishers and great blue herons as representative species, indicates that piscivorous bird populations are likely at negligible or low risk in the PSA. Further, even if ospreys were retained for evaluation, correction of the flaws in the osprey HQs would show negligible risks to migrating ospreys in the PSA.*

Piscivorous Mammals (Charge Question 3.7; Comments Section 10)

The ERA concludes that there are intermediate to high risks to mink and river otter due to PCBs and TEQs in the PSA, as well as downstream of the PSA. This conclusion is based on a number of data interpretations and assumptions that are not well supported by the evidence.

EPA's mink feeding study. In this study, farm-raised adult female minks and their offspring were fed a diet containing fish from the PSA at five PCB dose levels up to 3.7 mg/kg, and effects on survival, reproduction, and development were evaluated. This study did not show effects on many of the endpoints evaluated (e.g., adult food intake rate, adult weight, breeding success, whelping success, litter size, organ histopathology). However, the study did report an effect on kit survival at 6 weeks at the highest dose level, and the ERA considers that level a lowest observed adverse effect level (LOAEL). Even accepting that finding, this study demonstrates that mink are less sensitive to the mixture of PCBs in Housatonic River fish than they are to PCB mixtures present at other sites, where more significant effects have been found in mink at lower doses. Moreover, the data from this study do not support the conclusion that PCB exposure had a significant effect on kit mortality at 6 weeks. Kit survival at that point was highly variable among all dose groups, and without data at higher doses, it cannot be determined whether the mortality at the highest dose used was simply more of that non-PCB-related variability. Further, no necropsy data are reported for the kits that died prior to 6 weeks, and necropsies on kits that died later indicated that their deaths were not due to PCB toxicity but to infections common in captive mink. It is therefore inappropriate to attribute the kit mortality prior to 6 weeks to PCB exposure. Finally, GE's statistical reanalysis of the data showed *no significant effect on kit survival at the 3.7 mg/kg dose level*. Hence, this dose level should be considered an unbounded no observed adverse effect level (NOAEL), rather than a LOAEL.

<u>EPA and GE mink and otter field surveys.</u> Both EPA and GE conducted observational field surveys of mink and river otter tracks and signs over several years. These surveys indicate the presence of mink and otter in the PSA. Although the ERA claims that the EPA survey showed less frequent observations of mink and otter in the PSA than in reference areas or than would be expected, the survey does not support that conclusion because it was too sporadic and did not account for habitat differences between the PSA and the reference areas (lakes and pond). In contrast, GE's survey, which was conducted over a longer period, found evidence of *significant utilization of the PSA by mink and otter*, indicating that several mink and otter used the area regularly as part of their home ranges. The ERA's attempt to undermine this survey on the ground that the majority of sightings occurred during snow tracking in winter is unwarranted, since snow cover provides a much more extensive area for tracking these mammals than the limited artificial scent posts used in snow-free months, and hence a greater tracking effort was made in winter.

<u>HQ analyses based on comparisons to the literature.</u> The ERA has based its HQs on literature-derived effects curves and thresholds. This is inappropriate since site-specific effects data are available from the mink feeding study (i.e., an unbounded NOAEL of 3.7 mg/kg PCBs in diet). Moreover, due to a number of errors in the literature-derived effects curves and thresholds, the resulting HQs are overly conservative, as shown by the fact that they also predict significant risks for reference areas.

Overall assessment. *The existing site-specific data do not demonstrate harm to mink and otter in the PSA. Based on the literature, there may be risks to mink and otter in the PSA due to consumption of PCB-containing fish at some exposure level. However, the ERA adopts a threshold that is too low for this site, and it overstates both the magnitude of the risks and the certainty of the conclusions.*

Omnivorous and Carnivorous Mammals (Charge Question 3.6; Comments Section 11)

The ERA assesses risks to omnivorous and carnivorous mammals based on field studies of small mammals, including GE's population demography study of short-tailed shrews, and model-based HQs for short-tailed shrews and red fox. For the short-tailed shrew, the ERA characterizes the risks in the PSA as intermediate, but uncertain. For the red fox, the ERA states in some places that there are intermediate, but uncertain, risks in the PSA and in other places that the risks are undetermined. In any case, GE has a number of concerns about the ERA's interpretation of the evidence.

EPA's small mammal trapping field study. This study showed no differences in the incidence of placental scars in small mammals across sites with a wide range of PCB tissue levels in the small mammals. Although this study has limitations, it provides no evidence of harm, rather than undetermined evidence, as stated in the ERA.

GE's shrew population demography field study. This study examined the demography of local populations of short-tailed shrews on floodplain grids with varying soil PCB concentrations. It found high densities of shrews in the PSA and no relationship between any demographic parameter measured (population density, survival, sex ratio, reproduction, growth, and body mass) and PCB concentrations in soil. It also showed that values for these parameters were within ranges reported in the literature. The ERA states that EPA has reanalyzed the data from this study using different average PCB soil concentrations for the grids, and has found a weak but significant relationship between those concentrations and survival of the shrews. However, a further reanalysis of the data by the study author using the same PCB concentrations estimated by EPA and the same statistical technique used by EPA *failed to confirm any significant relationship between PCB concentrations and shrew survival*. Thus, this study provides no evidence of harm, rather than undetermined evidence, as stated in the ERA.

Modeled exposures and effects. The ERA's model-based HQ analyses for both shrews and red foxes used generic modeled food intake rates for mammals when measured species-specific rates are available, and used effects metrics based on rodents. As a result, the HQs are highly uncertain and conservative.

Overall assessment. *Given the strength of the field data on shrews and the uncertainty and conservatism in the model-based HQs, the data support a conclusion of low or negligible risks to shrews in the PSA.*

For red fox, as the ERA recognizes in some places, the risks are undetermined, because the only line of available evidence of harm (the HQs) is highly uncertain, and thus there are insufficient data to draw defensible conclusions about the presence and magnitude of harm to red fox in the PSA.

Threatened and Endangered Species (Charge Question 3.8; Comments Section 12)

The ERA assesses risks to T&E species based solely on modeled exposures and effects (i.e., HQs) for three species – bald eagles, American bitterns, and small-footed myotis. It concludes that there are high risks to bald eagles and American bitterns due to PCBs in the PSA and undetermined risks to small-footed myotis. However, the HQs substantially overestimate risks for these species due to the inputs used.

Modeled exposures and effects for bald eagles. The HQs for bald eagles are unnecessarily conservative and uncertain because they rely on: (a) a modeled generic food intake rate based on birds in general, when measured food intake rates for free-living bald eagles are available in the literature; and (b) PCB effects metrics for other species (chickens and kestrels) when a species-specific effects metric can be derived for PCBs from available high-quality bald eagle egg data. Correction of these deficiencies would reduce the bald eagle HQs for PCBs by 20-fold.

Modeled exposures and effects for American bitterns. The HQs for bitterns are likewise overly conservative because they rely on inappropriate PCB effects metrics (again based on chickens and kestrels) when more supportable effects metrics are available. Use of the latter would reduce the bittern HQs for PCBs by 12-fold.

Modeled exposures and effects for small-footed myotis. The HQs for small-footed myotis have substantial limitations and uncertainties due to use of exposure and effects assumptions based on other species (tree swallows and rats, respectively). Although GE has not identified solutions to these limitations, the limitations and uncertainties of these HQs need to be more fully recognized in the ERA.

Overall assessment. *The ERA overstates the risks to bald eagles and American bitterns because it fails to recognize the uncertainties associated with reliance on a single measurement endpoint (HQs) and uses needlessly over-conservative assumptions in the HQs. Correction of those assumptions would reduce the HQs by up to 20 times. However, the conclusions would still be highly uncertain due to use of only one measurement endpoint. The ERA correctly concludes that risks to small-footed myotis are undetermined.*

Extrapolation to reaches downstream of the PSA. The ERA concludes that there are risks to wintering bald eagles at Rising Pond. That conclusion is unwarranted because: (a) the ERA assumes that bald eagles would obtain 100% of their diet from Rising Pond when that impoundment makes up only about

1% of a wintering bald eagle's foraging range; and (b) the MATC used in this analysis was derived using a toxicity threshold for eagle eggs that fails to take account of a relevant study.

Discussions and Conclusions (Charge Question 4; Comments Section 13)

Section 12 of the ERA presents a summary of the ERA's findings and conclusions on specific receptors, an assessment of potential risks to biota downstream of the PSA and to species other than those evaluated specifically, a discussion of the "ecological implications" of the results, a summary of sources of uncertainty, and overall conclusions. There are several serious problems with Section 12 of the ERA.

Inconsistent or inappropriate weight-of-evidence evaluations. In many cases, the ERA gives *too low weight to site-specific field studies* that provide direct evidence of whether local populations or communities of receptors in the PSA are being adversely affected by PCBs. In other cases, the ERA asserts that these field studies provide *undetermined evidence of harm when in fact they showed no evidence of harm*. Further, where both EPA and GE experts performed field studies, the ERA *consistently gives greater weight to EPA's studies* over GE's studies, regardless of the quality of the study and the data and without adequate justification.

Overemphasis on HQs. Section 12 of the ERA *overemphasizes the importance of the HQs* by making comparisons across receptors based solely on the HQs, without taking into account the other lines of evidence or the differences in quality or certainty among the HQs. It also erroneously states that the HQs are not conservative since they generally did not use safety factors. In fact, there are numerous other sources of conservatism in the HQs, as discussed above.

Extrapolations to downstream reaches. The ERA's extrapolations of potential risks to reaches downstream of the PSA are *overstated* because they are based on overly conservative MATCs and, in the case of amphibians and bald eagles, fail to take account of habitat and foraging limitations downstream of the PSA.

Extrapolations to other species. The ERA's extrapolations of risks from the studied receptors to other species are *unduly speculative*, because they are: (a) unsupported by any data on those other species; (b) based only on exposure-related factors, thus failing to take account of inter-species differences in toxicological sensitivity to PCBs and TEQs; and (c) based on the highest predicted risk (i.e., HQs), not the overall risk conclusion, for each representative species.

Theoretical explanations for why abundant populations do not show a lack of PCB effects. In its "Ecological Implications" section, the ERA presents a number of theoretical factors that it asserts could explain how populations of receptors could be abundant at the site and still be suffering adverse effects of PCBs. It suggests that removal of predators could allow prey populations to remain abundant despite PCB effects, that immigration could be compensating for losses due to PCBs, and that PCB exposure may be making some populations more vulnerable to other stressors in the future even if no adverse effects are observed at present. *This attempt to explain away the findings of abundant populations in the PSA is entirely speculative, unsupported by any studies in the ERA, and has no place in the ERA.*

Overall conclusions. While the ERA includes many site-specific studies and multiple lines of evidence for many assessment endpoints, as well as use of a weight-of-evidence approach, it has many substantial flaws. These include: (a) overemphasis on individual-level rather than population-level effects: (b) underweighting of site-specific field studies that show healthy and thriving populations and diverse communities in the PSA; (c) overemphasis on HQs; and (d) incorrect interpretations or weighting of several studies. In fact, in most instances, in making judgments about the interpretation or weighting of particular lines of evidence, the ERA has adopted the interpretation or relative weighting that leads to a conclusion of risk, often stretching or misinterpreting the data to do so, and has downplayed the evidence indicating an absence of harm.

As a result, the ERA *substantially overestimates the risks of PCBs* to ecological receptors in the Rest of River area and cannot be considered objective, consistent, or reasonable (as specified in the Peer Review Charge) or accurate, reliable, or unbiased (as required by the EPA [2002a] Information Quality Guidelines). GE believes that a more balanced and supportable interpretation and evaluation of the evidence would result in the conclusions described above and summarized in Table ES-1 for each receptor group.

If the flaws in the ERA identified in these Comments are not corrected, GE believes that the ERA will not present an accurate characterization of ecological risks in the Rest of River area and cannot serve as a supportable basis for making a remedial action decision for this site.

Table ES-1

Summary of Conclusions for Receptor Groups

Receptor Group	Conclusion
1. Benthic Invertebrates	EPA's toxicity test data, properly interpreted, would support a sediment effects threshold of around 8 mg/kg PCBs. However, that threshold likely overestimates risks to benthic invertebrates, since there were no adverse PCB effects on benthic invertebrate community parameters in the community field study (at either fine-grained or coarse-grained sites), even at sites with higher PCB concentrations.
2. Amphibians	EPA's wood frog toxicity study provides some evidence of increased larval and metamorph abnormalities among amphibians in the PSA. However, the toxicity and field data both indicate that these abnormalities do not translate into adverse effects on local amphibian populations, and the field data reveal reproducing amphibian populations in the PSA.
3. Fish	EPA's toxicity studies showed no consistent evidence of adverse effects from PCB exposure, and the field studies showed self-sustaining fish populations and communities in the PSA. These data indicate that, while some effects may be occurring in individual fish, they do not affect the populations. Thus, there are negligible risks to the overall fish populations and communities in the Housatonic.
4. Insectivorous Birds	The strong field studies showing no effects on tree swallows and American robins, together with the correction of the HQ analyses, indicate that insectivorous birds are at negligible risk from exposure to PCBs and TEQs in the PSA.
5. Piscivorous Birds	Based on field data showing no effects on belted kingfishers or great blue herons, and considering the over-conservatism in the HQ analyses, piscivorous bird populations are likely at low or negligible risk in the PSA. Breeding ospreys should not be used as a representative species; correction of the osprey HQs to apply to migrating ospreys would show negligible risk to them.
6. Piscivorous Mammals	The mink feeding study (as reanalyzed) did not show adverse effects on mink at 3.7 mg/kg PCBs in diet, and the field survey data showed considerable usage of the PSA by minks and otter. Based on the literature, there may be risks to mink and otter in the PSA due to consumption of PCB-containing fish at some exposure level. However, the threshold must be > 3.7 mg/kg in fish, and any risk conclusions are uncertain given the lack of site-specific data showing harm to mink and otter.
7. Omnivorous and Carnivorous Mammals	The strong field study showing no effects on short-tailed shrews, coupled with the uncertainty and over-conservatism in the model-based HQs, support a conclusion of low or negligible risks to shrews in the PSA. For red fox, the risks are undetermined, because the only evidence of harm (the HQs) is highly uncertain, and there are thus insufficient data to draw defensible conclusions about harm to red fox in the PSA.
8. Threatened and Endangered Species	The only measurement endpoints – HQs – overestimate the risks to bald eagles and American bitterns because they rely on unnecessarily over-conservative assumptions. Correction of those assumptions would reduce the HQs by up to 20 times. However, any conclusions would still be highly uncertain due to use of only one measurement endpoint for each species. The risks to small-footed myotis, which are also assessed solely on the basis of highly uncertain HQ analyses, are undetermined.

1. INTRODUCTION

The General Electric Company (GE) submits these Comments to the U.S. Environmental Protection Agency (EPA) and the Peer Review Panel on EPA's public review draft of the *Ecological Risk Assessment for General Electric (GE)/Housatonic River Site, Rest of River* (hereinafter "ERA"), dated July 2003. The objective of these Comments is to provide EPA and the Peer Review Panel with additional information, viewpoints, and analyses relating to the ERA's characterization of potential risks to ecological receptors due to polychlorinated biphenyls (PCBs) and dioxin toxicity equivalents (TEQs) in the Rest of River portion of the Housatonic River and its floodplain.

The ERA integrates a number of underlying studies, including field and laboratory studies conducted by EPA contractors (frequently referred to herein as "EPA's studies"), as well as desktop modeling and literature-based analyses by EPA. The underlying studies also include numerous site-specific field studies conducted by GE contractors and/or academic experts retained by GE (frequently referred to herein as "GE's studies").

These Comments generally follow the structure of the Peer Review Charge (with relevant charge questions noted in the parentheses at the beginning of each comment). However, in addressing Charge Question 3, which relates to the eight assessment endpoints evaluated in the ERA, we first discuss a number of general points that apply to multiple assessment endpoints. We then discuss primary issues that are specific to each of the eight assessment endpoints. We are also providing several attachments that contain more detailed evaluations of some of the underlying EPA studies and more detailed descriptions of the GE-sponsored studies.[1] Some of these attachments were authored or co-authored by academic experts retained by GE. Finally, GE's views on each of the ten sub-questions within Question 3 are summarized for each assessment endpoint in tabular form at the end of the section addressing that endpoint.

The Peer Review Charge directs the peer review panel to evaluate the "objectivity, consistency, and reasonableness" of the various components of the ERA, both in applying EPA guidance and policy and apart from EPA guidance and policy. EPA's recent Information Quality Guidelines specify that "objectivity" includes "whether the disseminated information is being presented in an accurate, clear, complete, and unbiased manner, and as a matter of substance, is accurate, reliable, and unbiased" (EPA 2002a, p. 15; see also p. 22). In these Comments, we point out the major respects in which we believe that the ERA does not meet these tests of objectivity, consistency, and reasonableness. These include

[1] Full reports on most of the GE studies are also provided as part of the ERA on a compact disk entitled *Ecological Field Studies Conducted by General Electric for the Housatonic River Project.*

areas where we believe that the ERA incorrectly interprets the data, gives too much or too little weight to certain lines of evidence, or reaches conclusions that are not supported by the underlying data. We also offer recommendations for improvements to address these issues. For the convenience of the reader, the major points in each subsequent section of these Comments are summarized in bullets at the beginning of each section.

The text of these Comments focuses on major issues. Other specific errors and inconsistencies in the ERA that were identified during the course of our review are listed in Attachment A. (That attachment is not intended to be a comprehensive listing of all other specific errors and inconsistencies in the ERA, but simply those that we identified during our review and that are not discussed in the text or other attachments.)

2. CHARACTERIZATION OF ECOSYSTEM (QUESTION 1)

Key Points

- The ecological characterization that was conducted for the ERA provides substantial information on the ecosystem and the fish and wildlife populations at the site, particularly in the Primary Study Area (PSA), which extends from the confluence of the East and West Branches of the Housatonic River in Pittsfield to Woods Pond Dam in Lenox/Lee.

- Since PCBs have been present at elevated concentrations in the local ecosystem, particularly the PSA, since the 1930s, the local fish and wildlife populations have been exposed to elevated PCB levels for many generations. Despite this lengthy exposure period, the ecological surveys conducted for the ERA found abundant, diverse, and thriving fish and wildlife populations in the Housatonic River and its floodplain, including numerous species of invertebrates, fish, reptiles, amphibians, birds, and mammals.

- In general, the ERA adequately describes the ecological characteristics of the site and adequately applies the ecological characterization results in the problem formulation. However, in three respects, it does not provide an adequate characterization or does not adequately take account of the available information:

  Ø In assessing the risks to amphibians downstream of the PSA, the ERA does not provide necessary information on the locations of wetlands or vernal pools downstream of the PSA that could provide habitat for larval amphibians.

  Ø By selecting breeding ospreys as a representative species, the ERA does not take account of the ecological characterization's observations of ospreys on only six occasions, all corresponding to fall migration.

  Ø The ERA's characterization of risks to bald eagles wintering downstream of the PSA (specifically, at Rising Pond) does not account for the fact that that impoundment would make up only about 1% of a bald eagle's expected foraging range.

2. CHARACTERIZATION OF ECOSYSTEM (Question 1)

In general, the ERA's ecological characterization of the site, which is provided in Appendix A (Volume 3) of the ERA, is adequate and appropriately applied in the problem formulation for the ERA. That characterization properly focuses mainly on the PSA, which extends from the confluence of the East and West Branches of the Housatonic River in Pittsfield to Woods Pond Dam in Lenox/Lee, since the majority of PCBs and the highest PCB concentrations in the Rest of River area are found in the PSA. The ecological field surveys conducted for the ERA provide significant data to characterize the site ecosystem, as summarized in Section 2.1 below. In a few cases, however, the ERA fails to provide sufficient information or to adequately account for the results of the surveys, as discussed in Section 2.2 below. Overall, the PSA contains a diverse, abundant and healthy fish and wildlife population notwithstanding the decades-long presence of elevated levels of PCBs.

2.1 Overview of Site

Between 1932 and 1977, PCBs were used at the GE facility in Pittsfield and some were released to the East Branch of the Housatonic River via wastewater and stormwater systems and groundwater. As hydrophobic chemicals, the PCBs generally partitioned to the organic fraction of river sediments. During periodic flooding of the river, some sediments (and associated PCBs) were deposited on the floodplain. The highest concentrations of PCBs in the surface water, river sediments, and floodplain soil, as well as in their associated biota, are found upstream of Woods Pond Dam (QEA and BBL 2003). In this reach, elevated PCB concentrations in the floodplain soil are generally limited to areas within the approximate 10-year floodplain. Downstream of Woods Pond Dam, PCB concentrations in sediments, floodplain soil, and biota are all substantially lower than those in the PSA (QEA and BBL 2003).

The Housatonic River within the PSA is generally characterized as a shallow, relatively slow-moving water body, with abundant submerged vegetation. This reach of the river can be roughly divided at the Pittsfield wastewater treatment plant (WWTP) into two approximately equal sections with different flow and channel characteristics. Upstream of the WWTP, the river has a somewhat higher gradient. In that reach, sediments tend to be relatively coarse-grained and the organic carbon content tends to be relatively low. Downstream of the WWTP, sediments are finer and siltier, and tend to have higher organic carbon content. In this reach, due to the influence of the dam at Woods Pond, the velocity decreases, the river is more impounded, and there are numerous backwater and wetland areas, thus creating a very complex ecological habitat. Woods Pond is the first impoundment downstream of the GE facility and contains aquatic habitat characteristic of a standing water environment. These differences in habitat significantly affect both benthic invertebrate and fish communities in the PSA.

The majority of the floodplain within the PSA is undeveloped, with much of it (approximately 75%) consisting of publicly owned land (e.g., Housatonic River Valley State Wildlife Management Area). Hence, human use of the PSA is predominantly recreational, with limited areas of residential and agricultural properties. The ecological habitats within the PSA are very diverse and include large open-water communities surrounded by forest, emergent and forested wetlands, upland forests, and open habitat, such as fields, meadows, and grasslands. As described in the ERA, 18 natural ecological communities are present in the PSA (Vol. 1, p. 2-7; Vol. 3, pp. II-13, Table 1-2).

Downstream of the PSA, the Rest of River site extends southwards from Woods Pond Dam through western Massachusetts and Connecticut and encompasses areas where chemicals from the GE facility have migrated. The habitat is this area differs substantially from that within the PSA, with a noticeable decrease in wetlands and open-water areas along the river. The portions of the Rest of River site downstream of the PSA in Massachusetts consist of large areas of broad floodplain, with some portions currently used for agricultural purposes. In Connecticut, the river is characterized by a narrow floodplain and several large impoundments and varies in velocity and depth, with discrete areas of rapidly moving water that support stocked coldwater fish, such as trout.

As noted above, PCBs have been present in the local ecosystem for approximately 70 years. Hence, the snapshot of the ecosystem that is reflected in the ERA's ecological characterization represents a system that has been exposed to PCBs over many generations. Since PCB concentrations have been elevated in the PSA since the 1930s, large-scale or population-level impacts – if they exist – should have occurred before now and thus should be seen clearly in field surveys and studies. However, no such obvious effects were observed in the field surveys and studies conducted for the ERA. On the contrary, the Housatonic River and its floodplain contain abundant, diverse, and thriving fish and wildlife, as is clear from the ERA's ecological characterization. For example, surveys conducted for the ecological characterization documented the presence of numerous invertebrate, fish, reptile, amphibian, bird, and mammal species, as listed at right (Vol. 3, Section III, pp. 2-7 - 2-12, 3-5, 3-8, 3-9, 4-7, 4-11 - 4-14, 5-6, 6-7). Of these, threatened or endangered (T&E) or Special Concern species included: 15 plant species, 1 mussel species, 1 reptile species, 3 amphibian species, 8 bird species, and 1 mammal species (Vol. 3, Section III, Map 1-2, pp. 2-41, 4-31, 4-32, 5-20 - 5-25, 6-33).

Fish and Wildlife Species Documented in EPA's Ecological Characterization

- **12 mussel species**
- **38 dragonfly species**
- **41 fish species**
- **5 reptile species**
- **14 amphibian species**
- **139 avian species**
- **42 mammal species**

It should be noted that, in addition to PCBs, several other constituents were also identified as chemicals of potential concern (COPCs) in the ERA. These included polycyclic aromatic hydrocarbons (PAHs) and other semi-volatile organic compounds (SVOCs), dioxins and furans, several metals, and pesticides. Due to the historical use of the areas along the Housatonic River for paper and textiles mills, mining operations, and other industries, there are multiple sources of these constituents, and thus the presence of such COPCs in the Rest of River area may not be related to releases from the GE facility.

2.2 Characterization of Ecosystem in ERA

In general, the ERA adequately characterizes the ecosystem at the site, particularly within the PSA. In at least three respects, however, the ERA does not provide an adequate characterization of the ecosystem or does not take adequate account of the findings of the ecological characterization. These respects are identified below and discussed in more detail in subsequent sections of these Comments:

- First, as discussed in detail in Section 6.6 of these Comments, the ecological characterization does not provide information on the locations of wetlands or vernal pools downstream of the PSA. As a result, the maximum acceptable threshold concentration (MATC) developed for amphibians is incorrectly applied to the entire 100-year floodplain downstream of the PSA, regardless of whether that area includes suitable habitat for amphibian larvae (the life stage upon which the MATC is based). Information on the extent and location of wetlands and vernal pools downstream of the PSA is publicly available through the Massachusetts Geographic Information System (MassGIS) (www.state.ma.us/mgis) and should have been used in the assessment of risks to amphibians in these areas.

- Second, as discussed in detail in Section 9.2 of these Comments, the ecological characterization's finding that ospreys were only observed on six occasions, all of which corresponded to fall migration, is not reflected in the ERA, either in the selection of representative receptors or in the definition of exposure factor values (e.g., foraging time).

- Third, as discussed in detail in Section 12.5 of these Comments, the characterization of risks to bald eagles wintering downstream of the PSA (specifically, at Rising Pond) does not account for the very large foraging ranges required for this species in winter [i.e., 1,880 hectares (ha), per Griffin and Baskett (1985)] and the fact that Rising Pond would make up only about 1% of this range for a single bald eagle.

3. SELECTION OF ASSESSMENT AND MEASUREMENT ENDPOINTS (Question 2)

Key Points

- EPA guidance makes it clear that assessment endpoints in an ERA should address local populations and communities of biota, rather than individual organisms (with the possible exception of T&E species). While effects on such populations and communities can be extrapolated from studies on individuals or groups of individuals, the assessment endpoints should remain focused on the local populations and communities.

- Of the eight assessment endpoints selected for the ERA, two (for benthic invertebrates and amphibians) include references to local communities, but also include other endpoints (e.g., survival, reproduction) that appear to apply to individual organisms. The remaining six assessment endpoints relate to the survival, growth, and reproduction of individual organisms, rather than directly to population- or community-level effects.

- To be consistent with EPA guidance, these survival, growth, and reproduction endpoints should not be evaluated in isolation, but insofar as they affect the local populations and communities.

- For the most part, for all eight assessment endpoints, the ERA focuses and bases its conclusions primarily on individual-level effects data and does not adequately consider the implications for the local populations and communities. For example:

  - Ø The ERA generally does not attempt to extrapolate from individual-level effects to potential population-level consequences (with a few exceptions).

  - Ø The ERA generally assigns greater weight to the studies of individual-level effects than to site-specific data that directly address local population and communities.

  This is inconsistent with the proper focus of an ecological risk assessment and with EPA guidance.

- Most of the measurement endpoints selected in the ERA to evaluate the assessment endpoints are, in concept, appropriate types of studies and analyses, provided that their results are evaluated in terms of their relevance to local populations and communities. However, in many cases, the ERA does not provide such an evaluation. Further, in many cases, the ERA's interpretation or weighting of the specific studies and analyses and the conclusions drawn from them are unbalanced, inconsistent, or not well supported by the evidence.

3. SELECTION OF ASSESSMENT AND MEASUREMENT ENDPOINTS (Question 2)

3.1 Selection of Assessment Endpoints

EPA guidance provides that Superfund remedial actions should be designed not to protect organisms on an individual basis, but to protect local populations and communities of biota (EPA 1999, p. 3). Thus, the first management principle for conducting an ecological risk assessment is to provide a basis for selecting a response action "that will result in the recovery and/or maintenance of *healthy local populations/communities of ecological receptors* that are or should be present at or near the site" (EPA 1999, p. 3, emphasis added).[2] In line with this principle, EPA guidance specifies that risk assessors should select assessment endpoints that: "(1) are ecologically relevant to the site; i.e., important to sustaining the ecological structure and function of the local populations, communities, and habitats present at or near the site, and (2) include species that are exposed to and sensitive to site-related contaminants." It states further that such risk assessments "should use site-specific assessment endpoints that address chemical specific potential adverse effects to *local populations and communities* of plants and animals" (EPA 1999, p. 5, emphasis added). Similarly, EPA's Information Quality Guidelines note that ecological risk assessments should address "populations if applicable," and make clear that the instances in which they may address entities other than populations are when they address communities and ecosystems, not individual organisms (EPA 2002a, p. 22 & note 22).

EPA guidance also points out that effects on local populations and communities of biota can be extrapolated from effects on individuals or groups of individuals using a lines of evidence approach (EPA 1999, pp. 1, 3). However, it is important to recognize that, while such extrapolations can be made, the assessment endpoints themselves should remain focused on the local populations and communities, not on the individuals.

The eight assessment endpoints selected in the ERA are listed in Volume 1, p. ES-11 and Table 2.8-1 (pp. 2-62 - 2-64). The first two of these assessment endpoints – those relating to benthic invertebrates and amphibians – include explicit references to "community structure" and "community condition," respectively. However, it appears that the remaining endpoints listed for these receptors – i.e., "survival, growth, and reproduction" of benthic invertebrates and "survival, reproduction, development, and maturation" of amphibians – are intended to apply to individual organisms, rather than the local

---

[2] The EPA (1999) guidance also notes, as an exception to this rule, that threatened and endangered species may be evaluated on an individual basis. In concept, however, this focus is justified on the basis that, given the stressed nature of a T&E population, effects on individuals could impact the local population.

communities.[3] Moreover, the other six assessment endpoints in the ERA relate to “survival, growth, and reproduction” of individual organisms, rather than directly to population- or community-level effects.

The ERA justifies this focus on individual-level endpoints by suggesting that these individual-level parameters “are expected to be strong indicators of potential local population-level effects” (Vol. 1, p. 2-66). However, to be consistent with EPA (1999) guidance, the ERA must specifically address how these individual-level endpoints would translate into impacts on the local populations or communities. In other words, the survival, growth, and reproduction endpoints should not be evaluated in isolation, but insofar as they affect the local populations or communities. The ramifications to local populations are of greatest concern because the population is the smallest ecological unit that can be protected in a meaningful way (Suter et al. 1993). While individual organisms are transitory, populations persist and can be monitored on a human time scale.

Assessment endpoints that focus on survival, growth, and reproduction of individual receptors are appropriate <u>only</u> if they include a rigorous evaluation of any potential risk of individual-level effects to the sustainability of local populations and communities.

For the most part, however, the ERA does not adequately consider the population-level implications of the evidence. For example, it generally does not attempt to extrapolate from the individual-level effects to population-level consequences (except for its population model for woods frogs and, to some extent, for fish), and it typically assigns greater weight to the studies of individual-level effects than to the available site-specific data that directly address populations or communities. Furthermore, the ERA’s conclusions are generally based on the individual-level endpoints, without assessing their population- or community-level implications. As a result, the ERA’s conclusions are applicable primarily to individuals, rather than local populations or communities, which is inconsistent with EPA (1999) guidance.

The ERA’s statement of the assessment endpoints in terms of individual-level effects has also significantly and improperly influenced its weight-of-evidence evaluations. By focusing on individual-level effects, studies that show impacts to individuals are given high weight. This skews the weight-of-evidence evaluation. One of the key attributes of that evaluation is “degree of association,” which relates

[3] This is not entirely clear, since the statement of these two assessment endpoints in Attachment B to the Peer Review Charge – i.e., “[s]urvival, growth, reproduction and structure of the benthic invertebrate community” and “[r]eproductive success, development, maturation, and condition of the amphibian community” – suggests that all the endpoints listed for these receptors are intended to refer to the local communities. If so, then these assessment endpoints are appropriate in concept, provided that the evidence is evaluated in terms of how it relates to impacts on those communities. However, the statement of these endpoints in the ERA suggests that this is not the case.

to “the extent to which the measurement endpoint is representative of, correlated with, or applicable to the assessment endpoint” (Menzie et al. 1996, p. 2-68). Where the assessment endpoint is defined in terms of individual-level effects, the weight-of-evidence evaluation for measurement endpoints that are likewise based on individual-level effects assigns High or Moderate/High scores for degree of association. However, if the assessment endpoints had been based on population-level effects, as recommended by EPA (1999) guidance, then those measurement endpoints that are based on individual-level effects would have received lower scores for degree of association.

3.2 Selection of Measurement Endpoints

To provide lines of evidence to evaluate the assessment endpoints, the ERA has selected a variety of measurement endpoints (Vol. 1, pp. 2-62 - 2-64). For the most part, GE believes that these measurement endpoints are, in concept, appropriate types of studies and analyses, provided that, as discussed above, their results are evaluated in terms of their relevance to and implications for local populations and communities of the receptors.[4] However, in some cases, specific flaws in the design and implementation of the studies or analyses substantially limit the types of conclusions that should be drawn from them; and in many cases, we do not agree with the ERA’s interpretation of the results of the studies and analyses, particularly with respect to their treatment in the weight-of-evidence evaluations. For example, for measurement endpoints based on modeled exposures and effects – i.e., Hazard Quotients (HQs) – we have concerns regarding the assumptions employed in both the exposure characterization and in the development of effects metrics, as well as unbalanced application of the weight-of-evidence evaluations. These issues are discussed generally in Section 4 and specifically in the later sections on the individual endpoints (Sections 5 through 12).

[4] Exceptions to this general conclusion relate to the use of the following as measurement endpoints: sediment Toxicity Identity Evaluations; comparison of sediment and water concentrations to generic Sediment Quality Values (SQVs) and water quality criteria; EPA’s anecdotal observations of frogs; and modeled exposures and effects in ospreys. As discussed in Sections 5.2, 5.3.1, 6.1.3, and 9.2 of these Comments, these measurement endpoints are not appropriate for inclusion in the ERA.

4. GENERAL ISSUES RELATING TO EVALUATION OF ASSESSMENT ENDPOINTS (Question 3)

Key Points

- The ERA incorporates a substantial quantity of site-specific data and applies weight-of-evidence analyses in an effort to resolve conflicting lines of evidence. However, some of its interpretations of the data and applications of the weight-of-evidence approach are inconsistent and not supportable.

- When the weight-of-evidence approach is applied qualitatively, as in the ERA, there are many opportunities for subjective judgments. Because EPA has been the judge of both its own studies and GE's studies, as well as of the interpretation and weighting of all the lines of evidence, there is an inherent potential for bias in the weight-of-evidence evaluations.

- Field studies provide site-specific and species-specific data on the actual condition of a given population of receptors that have been exposed to PCBs through many generations. However, the ERA assigns unduly low weight to many of the field studies, particularly relative to the weight given to modeled exposures and effects.

- By contrast, the ERA places too much weight and emphasis on HQs based on modeled exposures and effects. These HQs are overly conservative because they:

  - Ø Do not address population-level impacts, but focus on individual-level impacts;

  - Ø Use inappropriate exposure assumptions for wildlife receptors, including: (a) modeled generic food intake rates when measured species-specific rates are available in the literature for most species evaluated; and (b) in some cases, unrealistic assumptions for foraging time within the PSA;

  - Ø Are mostly based on effects metrics derived from laboratory studies on other species and/or using different PCB mixtures;

  - Ø Use overestimated exposure point concentrations based on either the maximum detected concentration or the 95% upper confidence limit on the mean (95% UCL), calculated using a statistical procedure (Land's H-statistic) that overstates the upper bound on the mean; and

  - Ø Conflict in several cases with the results of field studies on actual populations of the receptors.

- The ERA also misapplies the weight-of-evidence approach by:

  - Ø Incorrectly defining evidence of harm as undetermined for several studies that in fact showed no evidence of harm; and

  - Ø Inappropriately assigning a magnitude of response to certain endpoints with undetermined evidence of harm.

- Some key studies and analyses are not adequately documented.

4. GENERAL ISSUES RELATING TO EVALUATION OF ASSESSMENT ENDPOINTS (Question 3)

This section discusses a number of general issues that affect the ERA's evaluation of multiple assessment endpoints and the overall approach employed to evaluate risk. It focuses primarily on the ERA's general evaluation of field studies and of HQs based on modeled exposures and effects, as well as its general application of the weight-of-evidence approach. GE's concerns relating to limitations in the design, implementation, and interpretation of the individual studies are detailed in the receptor-specific comments (Sections 5 through 12 of these Comments).

The ERA for the Housatonic Rest of River site is one of the most complex ecological risk assessments conducted to date anywhere in the country. One of the challenges of undertaking such a complex risk assessment, which involves numerous assessment endpoints and an even greater number of measurement endpoints, is the interpretation of the various, often conflicting, lines of evidence. Any given receptor may be evaluated through a combination of field studies, laboratory toxicity assays, and/or HQs based on modeled exposures and effects. These various lines of evidence inevitably differ with respect to findings and degree of scientific defensibility.

Weight-of-evidence evaluations attempt to assess the relative scientific defensibility of the various lines of evidence and balance the outcomes of those lines of evidence to yield a single unifying statement regarding the risks associated with a given assessment endpoint. EPA guidelines support this approach (e.g., EPA 2002b, p. 26). However, when this complex process is applied in a qualitative manner, as it is in the ERA, it can be quite subjective. In this case, EPA conducted the weight-of-evidence evaluations for both its own studies and GE's studies – i.e., it was the judge of both its own work and the work of others. As such, there is an inherent potential for bias. Moreover, in the interpretation or weighting of the various lines of evidence, there were many opportunities for interpreting the data or weighting the lines of evidence in a manner that favors a conclusion of risk. Thus, it is critical for the peer reviewers to closely examine the weights, evidence of harm, and magnitude of response reported for each measurement endpoint in the weight-of-evidence evaluations, as well as the conclusions reached for each receptor group, to evaluate whether those interpretations are objective, consistent, and reasonable (as specified in the Charge) and are accurate, reliable, and unbiased (as required by EPA [2002a] guidelines).

4.1 Evaluation of Field Studies

Numerous field studies were conducted as part of or in connection with the ERA, both by EPA contactors/experts and by GE contractors/experts. In general, advantages of such field studies include their site-specificity, species-specificity, and stressor-specificity, as well as their ability to reflect the actual exposure patterns and responses of the local population to those exposures. These advantages are magnified where, as here, multi-generational exposures have occurred over a period of decades. In particular, by studying receptors in the field, their responses to COPCs can be examined against the backdrop of natural population dynamics, reflecting the influences of density-dependence, predation, competition, and fluctuating climatic conditions. Because field studies often focus directly on a given assessment endpoint, the measurement endpoint and the assessment endpoint are essentially equivalent (particularly if assessment endpoints are appropriately defined in terms of the population).

FIELD STUDIES CONSIDERED IN ERA

- **Benthic Invertebrate Community Assessment (EPA)**
- **Amphibian Community Study (EPA)**
- **Wood Frog Early Life-Stage Study (GE)**
- **Leopard Frog Egg Mass Survey (GE)**
- **Fish Community Studies (EPA and GE)**
- **Largemouth Bass Reproduction/Population Structure Study (GE)**
- **Tree Swallow Nest Box Study (EPA)**
- **Robin Productivity Study (GE)**
- **Belted Kingfisher Productivity Study (GE)**
- **Mink and Otter Surveys (EPA and GE)**
- **Shrew Population Demography Study (GE)**

These studies provide site-specific and species-specific information on actual impacts on local populations of receptors that have been exposed to PCBs over many generations, taking into account natural population dynamics.

The many strengths of field studies are not adequately recognized in the ERA, due to inconsistent weight-of-evidence evaluations. In some cases, the ERA assigns field studies unduly low weights, particularly relative to the weights assigned to HQs based on modeled exposures and effects. For example, for fish and amphibians, field studies are assigned lower weights than EPA's comparisons of tissue data to toxicity levels derived from the literature; and for shrews, the HQs are assigned the same weight as the site-specific field study of shrew population demography. While the field studies have some limitations, they have a strong linkage to the appropriate assessment endpoint, and they are site-specific, spatially representative, and temporally representative. As such, they should generally be accorded greater weight than currently given in the ERA. These concerns are further discussed in subsequent sections of these Comments.

4.2 Evaluation of HQs Based on Modeled Exposures and Effects

The ERA uses modeled exposures and effects to calculate HQs for the majority of receptors, typically assigning them Moderate or Moderate/High weights. In addition, the ERA's Risk Summary (Section 12 of ERA) presents probabilistic HQs for all receptors "to facilitate comparison of risks among aquatic life and wildlife receptors and to give a broad overview of the findings of the risk assessment" (Vol. 2, p. 12-18). These practices place too much weight and emphasis on the HQs relative to other lines of evidence and obscure the high degree of conservatism and uncertainty associated with the HQs, as discussed below.

First, HQs do not address population-level impacts. Even probabilistic HQs actually only estimate exposures to either hypothetical individuals that are beyond the actual range of the population or to a very small number of individuals within the population. Because estimated doses are presented as distributions, there is the appearance that the HQs represent exposures across the entire population. However, for all receptors except mink and red fox, the HQs assume that the receptors derive 100% of their food from the PSA and only consume prey that contain the highest concentrations of COPCs. These assumptions restrict the applicability of the output to a small number of individuals (if any). The choice of effects metrics also limits the applicability of the HQs to individual-level responses. Because the effects metrics are based on individual-level rather than population-level responses, the comparison of estimated doses to those effects metrics only provides information on whether such individual-level responses are expected across the range of estimated doses. Neither the weight-of-evidence evaluations nor the uncertainty analyses adequately acknowledge the individual-level focus of HQs.

Second, the HQs for wildlife receptors were calculated based on highly conservative literature-derived exposure assumptions and modeling, rather than measured species-specific data. For example, for all such receptors, food intake rates were modeled using general algorithms that are not species-specific, but were developed for the overall taxonomic class or order to which the receptor belongs (e.g., birds, mammals). Such models require inputs regarding metabolic rate, food preferences, prey assimilation efficiencies, and gross energies – all of which have limited available data, yet strongly influence the results. However, for most wildlife receptor species evaluated, measured species-specific food intake rates are available in EPA's *Wildlife Exposure Factors Handbook* (EPA 1993), which states (p. 3-1) that such measured species-specific values are preferred over modeled intake rates based on allometric equations. In most cases, the ERA's modeled food intake rates are substantially higher than the values reported in the *Wildlife Exposure Factors Handbook* (EPA 1993) for free-living organisms. Without exception, the sensitivity analyses conducted as part of the ERA show that the input variables for the free metabolic rate (which in turn determines the food intake rate) are the most sensitive parameters of the entire HQ calculation.

Third, most of the effects metrics employed in the model-based HQs were derived from laboratory bioassays performed using different test species from the receptor species of interest and/or different mixtures of PCBs. For example, the HQs for robins are based on toxicity data for white leghorn chickens and American kestrels (Vol. 2, p. 7-74), despite the availability of site-specific toxicity data on both robins and tree swallows. Similarly, despite the availability of site-specific, stressor-specific effects data for mink, the effects curve used for the mink HQs (Vol. 2, p. 9-48) was derived from literature studies based on different mixtures of PCBs from the mixture in the Housatonic River. Thus, it is unclear whether the effects metrics used for these receptors are relevant to the actual responses of the species of interest under natural conditions following exposure to the mixture of PCBs present at this site.

Fourth, the Monte Carlo HQs calculated for wildlife receptors employed overly conservative exposure point concentrations (EPCs). These EPCs were based on the lower of the maximum detected concentration or the 95% UCL. In all cases, a statistical procedure known as Land's H-statistic was used to calculate the 95% UCL. This practice contradicts the decision criteria specified in Appendix C.5 of the ERA (Vol. 4, p. C.5-4), as well as EPA (2002b) guidance, both of which indicate that this method should *only* be used for data statistically determined to be lognormally distributed. In fact, approximately one-third of the data used in the wildlife assessments failed the statistical test applied for lognormality (Vol. 2, p. 6-11). Moreover, even for datasets that pass a statistical test for lognormality, it is widely recognized, including in EPA guidance (Singh et al. 1997; EPA 2002b), that Land's H-statistic can and often does substantially overestimate the actual upper bound on the mean if the dataset departs even slightly from a true lognormal population (even when it tests as lognormal on statistical tests) or when the sample size is small (i.e., $n < 30$), as it is for many of the datasets here. These points are presented in more detail in Attachment B to these Comments. As also discussed in Attachment B, GE believes that, instead of using Land's H-statistic to calculate the 95% UCLs, EPA should use Hall's bootstrap procedure (which is recommended in Appendix C.5 of the ERA as suitable for data that fit neither a normal nor a lognormal distribution) to calculate the 95% UCLs for use as EPCs in the wildlife HQs.

The over-conservatism of the HQs is evident from comparisons of predictions made for those endpoints evaluated based both on site-specific field studies and on HQs. For example, in the assessments of insectivorous birds (both tree swallows and American robins) and of belted kingfishers, the HQs predict high risks, while the site-specific field studies for these species all found no evidence of harm. Similarly, in the assessments of benthic invertebrates, the HQs predict intermediate to high risks at sites at which EPA's benthic invertebrate community study found no evidence of harm. Again, the total PCB (tPCB) HQ for short-tailed shrews predicts high risks, while the shrew population demography field study indicates no evidence of harm.[5] Hence, where both HQs and site-specific data are available, the HQs consistently indicate more risk than do site-specific data.

LIMITATIONS OF HQs BASED ON MODELED EXPOSURES AND EFFECTS

- **Focus on individual-level effects, not population-level effects.**
- **Use inappropriate exposure assumptions for wildlife – e.g., modeled generic food intake rates (not species-specific), unrealistic assumptions for foraging time in PSA.**
- **Mostly use effects metrics derived from laboratory studies on other species (and/or using different PCB mixtures).**
- **Use overestimated exposure point concentrations – maximum detected concentration or 95% UCL, calculated using Land's H-statistic, which substantially overstates upper bound where data are not lognormal or sample size is small (< 30).**
- **Conflict with results of site-specific field studies on local populations.**

Given the above factors, the Moderate or Moderate/High weights assigned to the HQ results in the ERA are generally unduly high, as will be detailed further in subsequent sections of these Comments.

4.3 Other Issues in Weight-of-Evidence Evaluations

In addition to the issues discussed above, we have identified two other ways in which the ERA appears to have misapplied the weight-of-evidence approach.

First, the weight-of-evidence evaluations for several endpoints conclude that evidence of harm (i.e., the outcome of the study) is undetermined, when in fact those studies showed no evidence of harm. For example, the ERA concludes that the GE and EPA fish field studies, the GE wood frog study and leopard frog egg mass survey, and the GE shrew population demography field study all yielded undetermined evidence of harm (Vol. 1, pp. 4-74, Table 5.4-5; Vol. 2, Table 10.5-3). In fact, whatever the limitations of

[5] The ERA inappropriately concludes that the shrew population demography study had undetermined evidence of harm, based on EPA's reanalysis of the data from that study. In fact, as shown in Section 11.2 and Attachment R of these Comments, even accepting the soil data used by EPA in that reanalysis, a revised statistical analysis continues to show no evidence of harm.

these studies, the studies showed no evidence of harm.[6] The ERA's failure to acknowledge that fact suggests a bias toward finding effects and downplaying the studies that did not show effects.

Second, for some endpoints assigned undetermined outcomes, the ERA assigns a magnitude of response. This is illogical: if an outcome cannot be determined, then its magnitude also must be undetermined. For example, the ERA assigns an undetermined outcome to some of the measurement endpoints evaluated for omnivorous and carnivorous mammals and to the HQ for small-footed myotis, but then designates their magnitude of response as intermediate or high (Vol. 2, Tables 10.5-3, 11.4-4). In these instances, the ERA inappropriately implies that an adverse effect was in fact shown by these measurement endpoints. On the other hand, it should be noted that the weight-of-evidence evaluations for benthic invertebrates, amphibians, and fish draw no conclusions regarding magnitude for those measurement endpoints with undetermined evidence of harm.

4.4 Incomplete Documentation of Key Studies

While the ERA provides many underlying studies and some underlying data, some key studies are not completely or adequately documented. This does not fully comport with the requirement in EPA's Information Quality Guidelines to present the information in a "comprehensive, informative, and understandable" way (EPA 2002a, p. 22). For example, while the vast majority of exposure data appears to be included on the ERA's Appendix L compact disk, the underlying site-specific data on effects are not included in that database, or anywhere else in the ERA. No raw data are included on the survival of mink kits, tree swallow clutch sizes, and field survey observations (data summarized only in Vol. 3, Appendix A.1). Without the underlying effects data, it is not possible to verify the accuracy of the overall analyses or conclusions. Moreover, the underlying effects data from EPA's fish toxicity studies are not clearly or completely presented in the ERA. In addition, in cases where the ERA reanalyzed GE's data (e.g., shrew population demography data, robin productivity data, belted kingfisher productivity data), insufficient information is provided to allow verification of the accuracy of those reanalyses. Finally, for endpoints that were tested for statistical significance, only the conclusions of the analyses are reported in the ERA (i.e., whether findings were significant or not significant). Degrees of freedom and the values of various statistics (e.g., t-statistic, F-statistic, p-value) should also be reported, in order to allow a more rigorous evaluation of the results.

---

[6] See note on page 4-6 above regarding the shrew population demography study.

5. BENTHIC INVERTEBRATES (QUESTION 3.1)

Key Points

- EPA's site-specific benthic invertebrate community study found no adverse effects of PCBs on the benthic community at sites with fine-grained sediments. While it did observe significant differences in benthic community parameters between PSA sites with coarse-grained sediment and reference sites, that observation does not show that PCBs are having an adverse effect at these sites because:

  - Ø The ERA did not adequately account for the many non-PCB-related habitat factors (e.g. substrate, water depth, flow, etc.) that may explain observed differences between PSA and reference sites;

  - Ø There is no correlation between PCB concentrations in sediment and those in the invertebrates' tissues at the coarse-grained sites;

  - Ø The study showed no significant exposure-response relationship between sediment tPCB concentrations at these sites (over a wide range) and any of the benthic community metrics; and

  - Ø An independent multiple regression analysis of the data shows that tPCB concentrations in sediments do not have a meaningful influence on benthic community metrics at these sites.

- EPA's site-specific toxicity tests do not support a sediment effects threshold of 3 mg/kg tPCBs.

  - Ø The method used in the ERA to calculate that threshold is flawed because it: (a) relied on sediment exposure PCB concentrations that are not representative of the sediments to which the organisms were actually exposed in the tests; (b) included an inappropriate comparison to the negative control; (c) used an averaging method that includes the double counting of certain endpoints; and (d) used 20% effects values that likely overestimate effects.

  - Ø A reanalysis of the toxicity test data, correcting for these errors, shows that those data support a sediment threshold of approximately 8 mg/kg tPCBs. Even this threshold, however, likely overestimates risks, as shown by the fact that the same stations used in the toxicity tests (which were fine-grained sites) showed no evidence of PCB effects in the benthic community study.

- The ERA's use of generic Sediment Quality Values and water quality criteria is inappropriate because those values do not reflect site-specific conditions.

- The ERA's literature-based threshold concentration of 3 mg/kg for tPCBs in invertebrate tissue is unsupported because it is not consistent with site-specific data and does not recognize the limitations of the underlying literature studies.

- **The ERA substantially overstates the magnitude and certainty of risks to benthic invertebrates in the PSA and in downstream reaches. The benthic community study showed no significant adverse PCB-related effects at either fine-grained or coarse-grained sites. While the toxicity test data, properly interpreted, would support a sediment effects threshold of 8 mg/kg tPCBs, that threshold likely overestimates risks, as shown by the absence of PCB effects in the community study at sites with higher concentrations.**

5. BENTHIC INVERTEBRATES (Question 3.1)

The ERA evaluates risks to benthic invertebrates based on three lines of evidence: (1) a site-specific benthic community field study; (2) site-specific toxicity tests and comparison of PCB concentrations in sediment to a sediment effects threshold derived from those tests; and (3) comparison of chemistry data for sediments, surface water, and benthic invertebrate tissue to benchmarks or effects levels derived from the literature. Based on these endpoints, the ERA concludes that there is a significant risk of adverse effects to benthic invertebrates throughout the PSA, with a moderate to high level of certainty. Potential risks are also predicted to occur in limited areas between Woods Pond and Rising Pond. These conclusions are based on a misinterpretation of the benthic community assessment, the toxicity test results, and the comparisons to literature-based benchmarks. Our chief concerns are discussed in the following subsections, supported by information in Attachments C and D. A summary of specific comments related to Questions 3.1(a)-(j) is provided in Table 5-2 (at the end of this section).

5.1 EPA Benthic Community Field Study

EPA contractors conducted a study of benthic community structure in 1999. Samples were collected from nine sites on the main branch of the Housatonic River and from four reference sites. Twelve synoptic samples were collected at each site for benthic invertebrate community data and sediment physical and chemical characterization. The ERA employs two basic strategies to evaluate the potential effects of PCBs on the benthic community: (1) a comparison of benthic community parameters at the PSA and reference sites; and (2) an evaluation of exposure-response relationships between tPCB concentrations in sediments and paired benthic community data. No significant differences were observed in benthic community parameters between the PSA sites with fine-grained sediments (located downstream of the Pittsfield wastewater treatment plant [WWTP]) and the fine-grained reference site, but the benthic community at PSA sites with coarse-grained sediments had lower abundance and richness and higher average rank plots (which integrate the results of multiple benthic metrics) than at coarse-grained reference sites. The ERA thus concludes that this study showed adverse effects of PCBs at the PSA sites with coarse-grained sediments, but not at the PSA sites with fine-grained sediments (Vol. 1, p. 3-62; Vol. 4, p. D-76).

There are several problems with this conclusion, which are summarized below and described in more detail in Attachment C to these comments, prepared by Dr. Scott Cooper (a benthic invertebrate expert from the University of California at Santa Barbara) and colleagues. First, although the ERA acknowledges that many habitat factors are known to affect the composition of benthic communities (e.g., Vol. 4, p. D-74), the EPA study evaluated only a few such factors (sediment grain size and total organic carbon) at the PSA sites, and not others that could impact the results (e.g., water depth, current speed, flow volume, light intensity, temperature). Moreover, the level of habitat analysis conducted at the reference sites was even less rigorous, raising questions about the comparability of habitat at the reference and PSA stations (Vol. 4, p. D-75). In addition, the ERA does not fully take account of the effect of grain size. As discussed in Attachment C (Section 3.1), an examination of sediment characteristics across the coarse-grained sites reveals the presence of nearly 100% sand at the three coarse-grained sites where the total abundance of invertebrates was significantly reduced (Stations 3, 4, and 5) (specifically, an average of 97.5% sand, compared to an average of 90.7% sand at the coarse-grained reference sites). Since it is well known that invertebrate abundances are low in areas dominated by sand (due to the unstable nature of these substrates), one would expect that, even in the absence of PCBs, the abundance of invertebrates at these three stations would be lower than at the coarse-grained reference sites.

Second, as also shown in Attachment C (Section 3.2), there is a lack of correlation between PCB concentrations in sediments and those in the tissues of the benthic organisms from coarse-grained sediments. These data indicate that that there is no relationship between the concentrations of tPCB in sediments and PCB exposure to the benthic community in coarse-grained sediments.

Third, although the stations with coarse-grained sediments had lower invertebrate abundance and richness and higher average rank plots than at reference stations, the ERA acknowledges that there were no significant relationships between the concentrations of PCBs in sediments (over a wide range of PCB exposures) and any of the benthic community metrics (Vol. 1, p. 3-58; Vol. 4, p. D-76). Given the absence of a significant exposure-response relationship and the lack of full habitat characterization at either the reference stations or the PSA stations, there is no basis for attributing the observed differences between the PSA stations and reference stations to PCBs, rather than to habitat or other uncontrolled variables.[7]

GE's experts have conducted supplemental multiple regression analyses on the data from the coarse-grained sites to evaluate relationships between tPCB, as well as certain other physical/chemical parameters, and benthic community structure at these sites. The physical/chemical parameters evaluated in these analyses included tPCB concentration, grain size, and total organic carbon (TOC) content; and

---

[7] In addition, although the study found no evidence of PCB-related effects on the benthic community in fine-grained sites, the ERA suggests that the lack of effects at these sites may be attributable to microhabitat variation, lower sediment concentrations, and reduced bioavailability of PCBs (Vol. 1, p. 3-62; Vol. 4, p. D-80). However, the data do not support these suggestions. The concentrations of tPCBs in most of the sediment samples from the fine-grained stations (e.g., Stations 6, 7 and 8) are well above the 3 mg/kg sediment-based effects threshold advocated in the ERA and the highest bioaccumulation of PCBs was observed in predators from Station 7, indicating that bioavailability is not limited at these stations (Vol. 1, Figures 3.2-3, 3.2-6; Vol. 4, Figures D.2-10, D.2-29) (Attachment C, Section 3.3).

the benthic community metrics included various indices of the benthic community.[8] These analyses are described in detail in Attachment C (Section 3.3). In summary, they show that the physical/chemical parameters evaluated were significantly correlated with some but not all of the benthic community metrics, and that even for those metrics, these parameters together accounted for only a relatively small portion of the variability in the metrics (up to around 30%). This indicates that other, unmeasured factors (e.g., other habitat variables) accounted for the majority of the variability in the metrics. Moreover, these analyses show that the tPCB concentration by itself accounted for only a very small fraction (1.0 to 6.8%) of the variability in these metrics. These results demonstrate that tPCB concentrations in sediments do not have a meaningful influence on benthic community structure, even at the coarse-grained sites.

In summary, based on the above analysis, it should be concluded that the benthic invertebrate community study found no significant adverse effects of PCBs at either the fine-grained sites or the coarse-grained sites. Given the fact that this study focused on actual impacts on the benthic community in the PSA, it should be entitled to high weight.

5.2 Site-Specific Toxicity Studies

Site-specific toxicity tests were conducted in 1999 by EPA contractors (Burton 2001, EVS 2003). Seven stations in the PSA were sampled for laboratory and in situ toxicity testing. Six of these sites were also used for the benthic community study. These stations were all from the area of the river downstream of the WWTP. As such, they represent the fine-grained sediments where no PCB-related effects were observed in the benthic community study. Laboratory testing included chronic toxicity tests evaluating growth, emergence, survival, and reproduction on *Hyallela azteca* (amphipod) and a *Chironomous tentans* (midge). *In situ* testing, with sediment and water only treatments, evaluated survival of the same organisms used in the laboratory bioassays as well as *Daphnia magna* (cladoceran) and *Lumbriculus variegatus* (oligochaete worm).

Based on these studies, the ERA develops a sediment effects threshold of 3 mg/kg tPCBs. The principal method used to do so is based on an evaluation of toxicity endpoints from the various bioassays (Vol. 1, p. 3-47; Vol. 4, p. D-48). Specifically, the ERA first calculates statistical endpoints, including 50% and 20% effects levels (i.e., LC50/IC50 and LC20/IC20 values), for each of the bioassays and acute and

[8] The ERA chose to concentrate its community analyses on simple measures of the absolute or relative abundances of invertebrates. Additional benthic community measures, such as diversity, evenness, and dominance, can provide a more complete understanding of benthic community structure and can show independent responses to environmental perturbations (Hurlbert 1971, Brower and Zar 1977, Pielou 1977, and Allan 1995). Therefore, in conducting this independent analysis, we included indexes of diversity and evenness, as well as measures of abundance.

chronic measured endpoints. It then calculates site-specific sediment MATCs based on the averages of the six lowest 50% and 20% effects levels calculated for the individual toxicity test endpoints (Vol. 1, p. 3-49; Vol. 4, p. D-53). These calculations are made for comparisons to the negative (laboratory) control, reference station A1, and reference station A3. The resulting average 50% and 20% effects levels calculated in the ERA for tPCBs are, respectively, 1.3 - 3.5 mg/kg and 0.1 - 0.9 mg/kg (Vol. 1, p. 3-49; Vol. 4, p. D-53). This calculation method is flawed for several reasons summarized below and discussed in more detail in Attachment D (Section 3).

(1) Inappropriate exposure point concentrations. The sediment concentrations employed in the ERA to calculate these effects levels are inappropriate. As the ERA recognizes, two sets of tPCB data were available for these calculations. The first consists of the results from sediment samples collected during both the laboratory and *in situ* bioassay studies specifically to characterize exposure and to provide a basis for evaluating bioassay-specific exposure-response relationships. The ERA refers to these as the "most synoptic" samples. The second consists of the "median" sediment tPCB concentrations from data collected during several sampling events between March and October of 1999, some of which are unrelated to the bioassays. Although the ERA calculates effects levels associated with both of these exposure estimates, it relies primarily on those calculated using the latter "median" concentrations to derive the effects threshold of 3 m/kg (Vol. 1, p. 3-49; Vol. 4, p. D-52). Those median concentrations, however, are derived from data collected at other times and thus are not representative of the actual concentrations to which the organisms were exposed in the toxicity tests.[9] The use of these non-synoptic data results in an overestimate of tPCB toxicity (see detailed discussion in Attachment D). For these reasons, effects thresholds should be based on the synoptic data collected explicitly for defining exposure in both the laboratory and in situ bioassays.

(2) Use of inappropriate comparisons to negative controls. The comparison to the negative control should not be included because, as acknowledged in the ERA, the controls did not account for "physiochemical factors that may mediate sediment toxicity" and because "differences in responses between negative control and reference sediment cannot be explained on the basis of PCB concentrations" (Vol. 4, pp. D-53, D-50).

---

[9] As an example, at Station 8A, a concentration of 31 mg/kg tPCBs was measured in the subsample of the homogenate used for the bioassay test. Because no *in situ* tests were conducted at this location, the remainder of the samples used to calculate the median came from a different study (samples collected in July and October). The resulting median value of 4.6 mg/kg tPCBs is thus not representative of the concentration to which the organisms were exposed in the laboratory and should not be used in the evaluation of effects in the bioassay test for this station.

(3) <u>Use of redundant endpoints.</u> No justification is provided for developing a threshold based on an average of the six lowest effects values (rather than the six distinct endpoints with the lowest levels). In several cases, the ERA includes values that are actually different measures of the same endpoint. This practice results in the double-counting of these endpoints. For example, the six most sensitive endpoints included in the ERA include two IC50s developed for *H. azteca* reproduction from the same test group, but at different times (i.e., 35-day and 42-day total number of young), as well as two LC20s for *H. azteca* survival at different times (35 and 42 days). These are simply different measures of the same basic endpoints. In addition, the weight endpoint is often double-counted by including endpoints for both dry weight and ash free dry weight, when these are both measures of the same effect (growth). (See Table D-2 in Attachment D for identification of all the redundant endpoints.)

(4) <u>Inappropriate reliance on 20% effects levels.</u> The use of 20% effects levels may overestimate effects if they cannot be statistically distinguished from the reference response. For the LC20 or IC20 values to be statistically distinguishable from the reference and thus valid, the minimum significant difference (MSD) between them must be less than or equal to 20%. Because the results of the ANOVA tests are not provided, it is not known whether the MSDs are greater than 20%; but because of the variability seen in the bioassay tests, it is likely that they are much higher. As a result, the most appropriate effects thresholds would be based on the 50% effects levels.

GE contractors have reanalyzed the data from the toxicity studies based on comparisons with References A1 and A3 (but not the laboratory control) using the most synoptic tPCB data and excluding redundant endpoints. This reanalysis is described in Attachment D (Section 4) and the results are summarized in Table 5-1 below. When this is done, the means of the lowest six 50% effects levels for tPCBs are, respectively, 8.3 mg/kg and 7.7 mg/kg. The means of the lowest six 20% effects levels are both 4.7 mg/kg tPCB. Because the relevance of the 20% effects levels is questionable (as noted above), GE believes that the most supportable toxicity threshold from EPA's toxicity tests would be based on the average of the 50% effect levels – i.e., between 7.7 and 8.3 mg/kg tPCBs. However, these values are likely overly conservative estimates of risks because of the questionable appropriateness of the reference sites and because other factors (e.g., other COPCs) could have contributed to the observed toxicity. This over-conservatism is demonstrated by the fact that the benthic community assessment showed no adverse

PCB-related effects at these same fine-grained stations, despite sediment tPCB levels considerably higher than 8 mg/kg (i.e., individual samples ranging up to approximately 50 mg/kg [Vol. 4, Figure D.2-10]).[10]

Table 5-1. Recalculation of Effects Thresholds
Using Synoptic Data and Excluding Redundant Endpoints

Species	Setting	Endpoint	LC50/IC50 (mg/kg tPCB)	Species	Setting	Endpoint	LC20/IC20 (mg/kg tPCB)
Comparison to Reference A1							
		Average	**3.5**			**Average**	**0.9**
C. tentans	lab	20-d AFDW	4.7	*D. magna*	in situ	48-h survival	1.3
D. magna	in situ	48-h survival	5.0	*C. tentans*	lab	20-d AFDW	1.9
H. azteca	in situ	48-h survival	8.3	*H. azteca*	lab	35-d survival	3.0
C. tentans	lab	20-d survival	<8.7	*H. azteca*	lab	42-d total young	4.5
C. tentans	lab	43-d emergence	<8.7	*C. tentans*	lab	20-d survival	<8.7
H. azteca	lab	42-d total young	14.4	*C. tentans*	lab	43-d emergence	<8.7
		Average	**8.3**			**Average**	**4.7**
Comparison to Reference A3							
C. tentans	lab	20-d AFDW	4.6	*C. tentans*	lab	20-d AFDW	2.0
H. azteca	lab	42-d total young	7.8	*H. azteca*	lab	35-d survival	2.0
D. magna	in situ	48-h survival	8.1	*H. azteca*	lab	42-d total young	3.3
H. azteca	in situ	48-h survival	8.2	*D. magna*	in situ	48-h survival	3.7
C. tentans	lab	20-d survival	<8.7	*C. tentans*	lab	20-d survival	<8.7
C. tentans	lab	43-d emergence	<8.7	*C. tentans*	lab	43-d emergence	<8.7
		Average	**7.7**			**Average**	**4.7**

Notes:

Based on "most synoptic" concentrations and six most sensitive endpoints; where there were multiple values for a single endpoint (e.g., survival at 35 and 42 days), the lowest value was used.

Used ash-free dry weight (AFDW) instead of dry weight.

Assumed 42-d total young, 35-d young, and 42-d young/female measured the same reproductive endpoint

In summary, while a more appropriate analysis of the toxicity test data would support a sediment effects threshold of 8 mg/kg tPCB, that threshold is most likely an overestimate of tPCB toxicity to actual benthic invertebrates in the PSA.[11]

---

[10] In addition to the method described above, the ERA uses a "general linear method" as a supplemental approach to evaluate overall trends in the data relating to exposure-response relationships (Vol. 1, p. 3-50; Vol. 4, p. D-54). This method is flawed because it also relies primarily on the non-synoptic "median" concentrations as exposure point concentrations (see Attachment D, Section 5). As a result, this method likewise overstates toxicity.

[11] In addition, the ERA gives Moderate weight to the results of the Toxicity Identification Evaluation (TIE) (Vol. 4, Table D.4-5). The ERA concludes from this Phase I TIE that tPCBs may be the main causal agent in the toxicity studies (Vol. 1, p. 3-52; Vol. 4, p. D-59). However, the TIE was Phase 1 and not a definitive study. It was not designed to provide an independent evaluation of effects, and the results of the TIE are inconclusive, showing only that non-polar organic chemicals were likely contributors to the observed toxicity. As a result, while the results of the TIE provide some evidence as to the degree to which observed effects can be linked to nonpolar organic chemicals rather than other COPC groups, the TIE study should be removed as a separate line of evidence in the weight-of-evidence analysis.

5.3 Comparison to Literature-Based Benchmarks or Effects Levels

5.3.1 Comparisons to generic sediment and water quality values

The ERA also relies on comparisons of sediment data to published Sediment Quality Values (SQVs) and comparison of water column PCB data to EPA's water quality criteria for PCBs (Vol. 1, pp. 3-63 - 3-66). SQVs are generic in nature and thus are primarily appropriate in screening-level ERAs when no site-specific data are available (see, e.g., Long et al. 1995). EPA's water quality criteria are likewise generic and thus may not reflect site-specific sediment-based effects. Given the availability of substantial site-specific data, the generic SQVs and water quality criteria should not be used to characterize risk in this ERA.

5.3.2 Derivation of tissue-based threshold concentration

The ERA developed a tissue-based effects threshold based on a qualitative examination of the distribution of effects and no effects data (Vol. 1, Fig. 3.3-14; Vol. 4, Fig. D.3-17) reported in literature studies (Vol. 4, Table D.3-9 & Att. D.1). The ERA identifies NOAELs and LOAELs based on a review of studies on uptake by and/or effects of PCBs on invertebrates. The ERA concludes that no adverse effects were observed at tissue residues below 3 mg/kg or less and that the frequency of adverse effects increased significantly at tissue residues above 10 mg/kg (Vol. 1, p. 3-54; Vol. 4, p. D-61). Therefore, it sets the tissue-based threshold at 3 mg/kg.

Data from the site-specific studies show that adverse effects on benthic invertebrates are unlikely to occur at the proposed threshold concentrations of 3 mg/kg in tissues. For example, the site-specific study on *Lumbriculus* reported in the ERA reported NOAELs at tissue concentrations of over 4 mg/kg (Vol. 1, Table 3.3-1; Vol. 4, p. D-44). Similarly, the benthic community assessment stations with fine-grained sediments had tissue residues up to 48 mg/kg, with no adverse effects (Vol. 1, p. 3-62, Figure 3.2-6; Vol. 4, p. D-76, Figure D.2-29). These site-specific data indicate that the literature-based tissue effects threshold is not a reliable predictor of effects in the field.

Furthermore, our review of the literature studies used as the basis for the tissue threshold (listed in Vol. 4, Table D.3-9 of the ERA) found that the quality of the data in these studies and the applicability of these data to predicting adverse effects to benthic invertebrates on the Housatonic River are questionable. The available data linking tissue concentrations with adverse effects are limited and inconsistent. Only five of the ten studies evaluated reported adverse effects (at concentrations ranging from 3.9 to 425 mg/kg). Of these five studies, four (Duke et al. 1970; Nimmo et al. 1974; Velduizen-Tsoerkan et al. 1991; Lowe et al. 1972) tested saltwater organisms, whose mechanism of response to tPCBs may be different from those of the potential receptors in the Housatonic River. As to the fifth study (Sanders and Chandler 1972), while it reported effects anecdotally, it did not evaluate whether the effects were statistically significant. Additionally, the endpoints measured in some of these studies, such as anoxia tolerance for bivalves (Velduizen-Tsoerkan et al. 1991) and shell growth for oysters (Lowe et al. 1972), may not be of ecological relevance to the protection of the benthic community in the Housatonic River. In short, the literature data are not sufficient for developing tissue-based effects levels for tPCBs in benthic invertebrates.

5.4 Overall Assessment

The ERA concludes that PCBs pose a significant risk of adverse effects on benthic invertebrates throughout the PSA, with a moderate to high level of certainty. GE believes that this is a substantial overestimate of both risk and certainty. A close evaluation of the benthic community data shows that there were no significant relationships between the benthic community parameters and tPCB concentrations in sediments. In addition, in calculating a site-specific sediment effects threshold of 3 mg/kg tPCBs from the toxicity bioassay data, the ERA uses inappropriate exposure data and calculation methods, which result in an underestimate of the likely threshold concentration. Further, the literature-based tissue effects threshold must be considered unreliable due to: (1) the lack of concordance between the concentrations at which effects are predicted and the site-specific effects data; and (2) the limitations of the underlying literature. Consequently, the HQs that depend upon these thresholds significantly overstate risk. The lack of any PCB-related effects on the benthic community at significantly higher PCB concentrations than estimated based on either site-specific or literature-based thresholds provides further support of the inappropriateness of those effects thresholds.

The ERA's weight-of-evidence evaluation (Vol. 1, Table 3.4-1; Vol. 4, Sec. D.4.5.1.1) is also internally inconsistent and flawed. It places too little weight on the benthic community study, which clearly demonstrates the lack of PCB-related effects over a wide range of sediment and tissue tPCB concentrations. For example, the benthic community endpoint is closely linked to the assessment endpoint of community structure but receives only a Low/Moderate weighting for the degree of biological association (Vol. 1, Table 3.4-1; Vol. 4, p. D-95). In addition, for stressor/response, the benthic community assessment receives a Low/Moderate weight because the endpoint "may respond to a broad range of stressors other than the COCs" (Vol. 4, p. D-95). However, the same is true for the sediment toxicity tests, which received a ranking of Moderate/High for this attribute. Moreover, there is no justification for the site-specific benthic community study to receive the same weight as (rather than a higher weight than) the literature-based tissue thresholds and the toxicity identification evaluation (TIE).

SUMMARY OF CONCLUSIONS
<u>RISKS TO BENTHIC INVERTEBRATES IN PSA</u>

The ERA overstates the magnitude of risks to benthic invertebrates in the PSA and the certainty of the conclusions.

- **The benthic community study showed no significant adverse PCB-related effects on benthic invertebrate community parameters at either fine-grained or coarse-grained sites.**

- **EPA's toxicity test data, properly interpreted, would support a sediment effects threshold of around 8 mg/kg tPCBs. However, that threshold likely overestimates risks to benthic invertebrates, as indicated by the absence of PCB-related effects in the benthic community study, even at sites with considerably higher tPCB concentrations.**

5.5 Extrapolation to Downstream Reaches

The uncertainties discussed above with regard to the PSA (e.g., lack of concordance between the benthic community and toxicity test results, and inappropriate derivations of the sediment and tissue thresholds) are amplified by the ERA's attempt to extrapolate risk to downstream reaches. The ERA identifies depositional areas in Rising Pond as potentially at risk. However, these depositional areas tend to be fine-grained and there was no evidence of effects in fine-grained sediment in the benthic community study. Furthermore, our detailed reanalysis of the data indicates that the observed patterns in community metrics are not closely related to sediment PCB concentration in coarse-grained sediments. These analyses do not support conclusions regarding potential risk to the downstream reaches.

See Table 5-2 for specific comments on Charge Questions 3.1(a)-(j)

Table 5-2. Assessment Endpoint: Community Structure, Survival, Growth, and Reproduction of Benthic Invertebrates

EPA Charge Question #3.1	GE Response
(a) Were the EPA studies and analyses performed (e.g., field study, site-specific toxicity study, comparison of exposure and effects) appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), and based on accepted scientific practices?	**Benthic Community Study:** Methods were generally sound. However, habitat characterization was inadequate to assure that reference sites were properly matched to PSA sites, thus confounding comparisons to reference sites. (Section 5.1) **Laboratory and In Situ Bioassays:** Methods employed in bioassays generally appear sound. However, methods used to develop toxicity thresholds from these data are flawed, resulting in overly conservative thresholds. (Section 5.2) **Comparisons to Sediment Quality Values and Water Quality Criteria:** Use of generic values is inappropriate, given the abundance of site-specific data for evaluating risk to benthic invertebrates. (Section 5.3.1) **Comparisons to Literature-Based Thresholds:** The analyses employed in developing literature-based tissue thresholds are flawed, resulting in overly conservative thresholds that are not consistent with site-specific data. (Section 5.3.2)
(b) Were the GE studies and analyses performed outside of the framework of the ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), based on accepted scientific practices, and incorporated appropriately in the ERA?	Not applicable.
(c) Were the estimates of exposure appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?	**Benthic Community Studies:** Exposure is appropriately quantified. **Laboratory and In Situ Bioassays:** Use of median values, which include data unrelated to the bioassays, rather than "synoptic data" collected in conjunction with the bioassays, is inappropriate and results in overly conservative estimates of toxicity thresholds. (Section 5.2)
(d) Were the effects metrics that were identified and used appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable)?	**Site-Specific Effects Metric:** The 3 mg/kg effects metric developed from the bioassays is inappropriate due to: (1) use of exposure data from samples unrelated to the bioassays; (2) inappropriate comparisons to a laboratory control; (3) use of redundant endpoints; and (4) inclusion of 20% effects levels that likely overestimate effects. A reanalysis of these data supports a threshold of 8 mg/kg. However, even this threshold is likely too low, as shown by the fact that the same stations used in the bioassays showed no PCB effects in the benthic community study, even at much higher PCB concentrations (individual samples up to 50 mg/kg). (Section 5.2) **Literature-Based Effects Metric:** The literature-based tissue threshold of 3 mg/kg tPCB is unsupported because it does not take account of the limitations of the underlying studies and is not consistent with the site-specific data, which showed no PCB effects on the benthic community at stations where invertebrates' tissue PCB levels were considerably higher than 3 mg/kg. (Section 5.3.2)
(e) Were the statistical techniques used clearly described, appropriate (i.e., objective, consistent, and reasonable), and properly applied for the objectives of the analysis?	**Benthic Community Study:** Statistical methods were generally appropriate. **Laboratory and In Situ Bioassays:** No statistical methods were employed to determine if 20% effects levels were significantly different from the reference results. In the absence of this determination, it is inappropriate to use the 20% effects levels in developing thresholds. (Section 5.2)

Table 5-2. Assessment Endpoint: Community Structure, Survival, Growth, and Reproduction of Benthic Invertebrates

EPA Charge Question #3.1	GE Response
(f) Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable)?	**Benthic Community Studies:** The discussion of the benthic community data lacks objectivity, because it fails to recognize the absence of a significant PCB impact on the benthic community at the coarse-grained stations (as well as the fine-grained stations). (Section 5.1) **Laboratory and In Situ Bioassays:** The bioassay data do not support sediment PCB effects threshold of 3 mg/kg for the reasons given in the response to Question 3.1(d) above. (Section 5.2) **Comparisons to Sediment Quality Values and Water Quality Criteria:** Use of these generic values is not appropriate, since site-specific data are available. (Section 5.3.1) **Comparisons to Literature-Based Tissue Thresholds:** The literature-based tissue threshold of 3 mg/kg tPCBs is flawed for the reasons given in the response to Question 3.1(d) above. (Section 5.3.2)
(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements could be made.	**All Studies:** Sources of uncertainties are generally identified.
(h) Was the weight of evidence analysis appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable)? If not, how could it be improved?	**Benthic Community Study:** The benthic community study is given too little weight (Moderate), since it directly evaluated potential effects on the benthic community in the PSA. (Section 5.4). **Laboratory and In Situ Bioassays:** Weight given to these bioassays (Moderate/High) is appropriate, but the data are not properly interpreted. (Section 5.2) **Comparisons to Literature-Based Tissue Thresholds:** Weight given to the comparisons to these thresholds (Moderate) is too high given the uncertainties in the underlying data and the existence of site-specific information. (Section 5.4) **Comparisons to Sediment Quality Values and Water Quality Criteria:** Should not be used as a separate line of evidence given the site-specific data. (Section 5.3.1) **Toxicity Identification Evaluation:** Should not be used as a separate line of evidence. (Section 5.2, note 11)
(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?	Predictions of risks to reaches downstream of the PSA are not supported because they (1) are based on a sediment MATC (3 mg/kg) that is too low and (2) apply to depositional areas with generally fine-grained sediments, when the benthic community study showed no evidence of harm at fine-grained stations with PCB concentrations substantially above 3 mg/kg. (Section 5.5)
(j) In the Panel members' opinion, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?	**The ERA overstates the magnitude of risks to benthic invertebrates in the PSA and the certainty of the conclusions.** The benthic community study showed no significant adverse PCB-related effects on the benthic community at either fine-grained or coarse-grained sites. While the bioassay data, properly interpreted, would support a sediment effects threshold of 8 mg/kg PCBs, that threshold likely overestimates risks, as shown by the absence of PCB effects in the benthic community study, even at sites with considerably higher PCB concentrations. (Section 5.4)

6. AMPHIBIANS (Question 3.2)

Key Points

- Contrary to the ERA's assertions, EPA's leopard frog toxicity study does not provide reliable evidence of impaired reproduction and development in leopard frogs in the PSA, because:

  - Ø The ERA relies heavily on comparisons between frogs from the PSA and frogs purchased from a commercial distributor, which were collected in very different locations and were subjected to less environmental and handling stress (including much shorter holding times); and

  - Ø The adult females from the PSA, which had low rates of mature eggs, may have been collected outside the prime reproductive season, confounding the assessment of both male and female fertility.

- In contrast to the EPA leopard frog study, GE's leopard frog egg mass survey showed significant reproductive activity in the leopard frog population in the PSA.

- The ERA's interpretation of EPA's site-specific wood frog toxicity study overemphasizes endpoints (i.e., malformation rates and sex ratio data) that are not directly relevant to the sustainability of the local population, while failing to consider endpoints that are more directly relevant to the local population such as survival, metamorphosis, and reproductive output (which showed no PCB effects).

  - Ø An independent analysis shows that there was no significant negative relationship between PCB exposure and net (abnormality-free) metamorph output in these data.

  - Ø The data indicate that adult wood frogs do not have skewed sex ratios and that tissue PCB concentrations were not associated with skewed sex ratios in metamorphs.

- EPA's wood frog population model is flawed due to its failure to consider density-dependence. As a result, it greatly overstates the risk of extinction for both PCB-exposed and unexposed wood frog populations – results that are refuted by the presence of woods frogs populations in the PSA.

- The ERA's site-specific sediment/soil MATC of 3 mg/kg tPCBs and tissue MATC of 1 mg/kg tPCBs are based primarily on the sex ratio data from EPA's wood frog study, which do not provide a reasonable basis for establishing effects thresholds. The EPA malformation data would support MATCs of 38 mg/kg in sediment/soil and 6 mg/kg in tissue, but these thresholds likely overstate population risks.

- The ERA's literature-based tissue effects threshold of 1 mg/kg tPCBs is based on a misinterpretation of one of the underlying studies and is considerably lower than that supported by the data.

- **Overall, while there is some evidence of increased larval malformations and metamorph abnormalities among amphibians in the PSA, the available data do not indicate that these malformations/abnormalities translate into adverse impacts on local amphibian populations.**

- The ERA overestimates risks to reaches downstream of the PSA by relying on an overly conservative sediment/soil MATC and failing to taking account of the limited amphibian habitat in those reaches.

6. AMPHIBIANS (Question 3.2)

The ERA evaluates risks to amphibian populations and communities based on three types of measurement endpoints: (1) site-specific toxicity studies conducted by EPA and GE contractors; (2) field surveys conducted by EPA and GE contractors; and (3) comparisons of site-specific concentrations of PCBs in frog tissues to tissue-based effects thresholds derived from the literature. Based on these endpoints, the ERA concludes that amphibians have a high risk of ecologically significant effects throughout the PSA, with a moderate to high level of certainty. It also develops a sediment/soil MATC of 3 mg/kg tPCBs and a tissue MATC of 1 mg/kg tPCBs, and its uses those MATCs to predict risks to amphibians breeding in floodplain areas downstream of the PSA.

The ERA substantially overstates PCB-related risks to amphibian populations and communities. Although there is some evidence of increased larval malformations and metamorph abnormalities, the levels of malformation appear to be insufficient to result in effects on the output of metamorphs and thus on the local populations. Most of our comments relate to the site-specific studies performed for EPA and GE; additional details concerning these studies are provided in Attachments E through H. A summary of specific comments related to Questions 3.2(a)-(j) of the Peer Review Charge are provided in Table 6-1 (at the end of this section)

6.1 Site-Specific Studies on Leopard Frogs

The ERA assesses risks to Northern leopard frogs (*Rana pipiens*) using EPA's site-specific toxicity tests, GE's leopard frog egg mass survey, and EPA's anecdotal observations. As detailed below, there are significant shortcomings and limitations in EPA's studies. In any event, one of the key conclusions drawn from those studies, that there is impaired reproduction of leopard frogs in the PSA, is contradicted by GE's egg mass survey, which showed that, in fact, leopard frogs are reproducing in the field under natural conditions.

6.1.1 EPA's leopard frog study

The objective of EPA's leopard frog study was to assess the effects of COPCs on reproduction, early development, and maturation (metamorphosis) in adults and larvae. The study was conducted for EPA in 2001 by Fort Environmental Laboratories (FEL) and Woodlot Alternatives (FEL 2002a). The original study design called for collecting adult males and females from nine sites in the PSA (target frogs) and from three reference areas. However, no adult leopard frogs were found at the reference sites (ERA, Vol. 1, p. 4-34). Therefore, reproductively mature leopard frogs purchased from a commercial supplier, Carolina Biological Supply (CBS), were used as so-called "external references." Moreover, most of the adult females collected from the PSA had immature oocytes, leading EPA to conclude that their reproduction was impaired. As a result, an additional field effort was undertaken to collect fertilized egg masses and larval frogs from the PSA for assessment of developmental effects. Again, no egg masses or larvae were found at the reference sites, and hence the offspring of the commercially purchased frogs were used for comparison to those from the PSA egg masses and larvae (Vol. 1, p. 4-39). Due to a number of limitations, which are summarized below and discussed in more detail in Attachment E, the leopard frog study does not provide an appropriate basis for evaluating PCB-related risks to amphibians.

First, the ERA relies to a large extent on comparisons between the target and the commercially purchased frogs to assess effects relating to the reproductive condition of adults and to larval growth and development. The commercially purchased frogs were collected from Vermont (ERA, Vol. 1, p. 4-39) and southern Canada (FEL 2002a, p. 24) (precise locations not provided). The use of these frogs is inappropriate because of differences in the origin and handling conditions of the target and commercially purchased frogs. Due to the differences in origin, it cannot be assumed that the reproductive status of the target and commercially purchased frogs was comparable at the time of collection. In fact, the authors of the underlying report have acknowledged that "*adult external reference specimens were not exposed to the same environmental stressors that the adult specimens collected from within the lower Housatonic River watershed were exposed*" (FEL 2002a, p. 28, emphasis added), and thus have excluded the data from the commercially purchased frogs from their statistical analyses (FEL 2002a, p. 19). Further, comparisons of "randomly" collected field specimens to hand-picked specimens provided by a commercial distributor, whose objective was to provide healthy, gravid female frogs and healthy mature male frogs, violates key assumptions inherent in any scientific study – i.e., that samples are selected using uniform criteria and methodologies (e.g., EPA 1996a,b).

In addition, the target and commercially purchased frogs experienced different handling stresses. Adult target frogs were held for up to one month before females were given hormone injections to induce ovulation (FEL 2002a, p. 45),[12] whereas the commercially purchased frogs were injected with hormones within 48 hours of transfer to the laboratory. The egg masses from the commercially purchased frogs were laid and larvae hatched in the laboratory and held until the study was initiated, whereas target egg masses and larvae were transported from the field during a sensitive developmental stage. Finally, if

12 The date of the attempted artificial fertilization was provided in Attachment D to an earlier draft of this report. It indicated that super-ovulation and artificial fertilization of eggs produced by target frogs from four of five PSA sites (W-9a, EW-3, E-1 and W-1) were not attempted until May 1 or 2, 2000, more than four weeks after most of the frogs were collected (between March 25 and April 22, 2000). No information regarding the dates artificial fertilization was attempted are provided in the final version of the report.

target frogs had mature oocytes when collected, extended holding times may have contributed to the failure of the hormone injections.

In addition, the ERA draws a number of unsupported conclusions regarding the "impaired" reproductive condition of adult frogs from the PSA. For example, the ERA relies on the fact that the target adult females had very low rates of mature eggs (Vol. 1, p. 4-35; Vol. 5, p. E-65). However, this could have resulted from poor timing in the sample collections. For example, FEL's adult frog collections occurred between March 25 and April 22, 2000, when surface water temperatures were ~8 to 10°C (Vol. 5, p. E-16) – at or below the temperature of 10°C at which leopard frog breeding is reported to begin (Wright 1920; Hine et al. 1981; Kendall 2001, 2002). Hence, these collections may have been too early to capture a representative sample of reproductively mature females. Further, FEL's egg mass collections occurred 2-4 weeks later, between May 8 and 23, 2000, which may have been too late to collect a large number of egg masses (see discussion in Section 6.1.2 below).

The ERA also concludes that male frogs from the PSA exhibited a trend of increasing sperm head abnormalities with increasing sediment tPCB concentrations (Vol. 1, p. 4-35; Vol. 5, pp. E-64, E-65). However, that trend was not statistically significant (FEL 2002a, p. 48). Because male leopard frogs disperse into the floodplain when not breeding, their tissue body burdens would not be expected to be correlated with PCB concentrations in pond sediment. The relationship between sperm head abnormalities and male tissue body burdens, which is more biologically relevant, was not assessed. Moreover, the ERA concludes that these sperm head abnormalities may have contributed to poor fertilization success of target frogs (Vol. 1, p. 4-39; Vol. 5, p. E-65). However, the viability of sperm cannot be evaluated when the oocytes being fertilized are immature. Thus, no conclusions can be drawn regarding the viability of sperm based on this study.

The ERA further concludes that the study showed adverse developmental effects in the larvae from the PSA (i.e., high mortality, increased malformations, developmental delays, and low incidence of metamorphosis) (Vol. 1, pp. 4-39 - 4-44; Vol. 5, pp. E-68 - E-73). However, to a large extent, these conclusions rely on comparisons to the larvae from the commercially purchased frogs, which, as noted above, cannot provide reliable evidence of PCB-related effects. Moreover, for some endpoints (mortality and incidence of metamorphosis), the data show no exposure-response relationships among the PSA larval groups (e.g., no differences between larvae from ponds with low sediment PCB concentrations and those from ponds with high sediment PCB concentrations) (Vol. 1, pp. 4-41, 4-42), thus indicating that some factor other than PCB exposure (e.g., transportation stress) was responsible for the high mortality and low incidence of metamorphosis in the PSA larvae. For the endpoints that do show differences between the PSA ponds with high and low sediment PCB concentrations (increased malformations and developmental delays), the results are limited by a lack of statistical analysis, small sample size, and the unclear biological relevance of these endpoints to the local leopard frog populations, given the lack of any difference between these ponds in the critical endpoints of survival and metamorphosis. (See Attachment E, Section 5.)

6.1.2 GE leopard frog egg mass survey

Given the findings of the FEL study, a leopard frog egg mass survey was conducted in 2003 by a GE contractor (ARCADIS) to determine whether adult leopard frogs are, in fact, suffering from reproductive dysfunction in the PSA under natural conditions, as claimed in the FEL study and the ERA. A summary of GE's survey is provided as Attachment F to these Comments.

Ponds containing potentially suitable breeding habitat for leopard frogs were identified before surface water temperatures reached 10°C (which, as noted above, is the temperature at which leopard frog breeding is reported to begin). This initial reconnaissance identified 44 ponds with habitat that might support leopard frogs. To be conservative, as many ponds as possible were kept in the study, including ponds with questionable habitat. The selected ponds represented a broad range of habitat characteristics and concentrations of tPCBs in sediment, both of which were representative of the overall spectrum of ponds in the floodplain. These ponds were then surveyed for leopard frog egg masses over 15 days during the breeding season.

During the survey, a total of 216 leopard frog egg masses were identified in 17 ponds throughout the PSA. In many instances, it was possible to confirm that these eggs had been fertilized because developing larvae were visible. These observations indicate that females from the PSA produce eggs that can be fertilized and that males produce sufficient viable sperm to fertilize them. There was no evidence of a relationship between concentrations of average tPCBs in the pond sediments and the number or incidence of leopard frog egg masses.

The striking difference between the number of egg masses found in this survey (n = 216) and the number found by FEL (2002a) (n = 6) suggests that some factor unrelated to PCB exposure (e.g., ponds sampled, timing of collection) most likely accounted for the low rates of mature eggs in adult females and low number of egg masses found by FEL. For instance, as mentioned above, FEL's adult frog collections (March 25 to April 22, 2000 per Appendix D of FEL 2002a, with surface water temperatures of 8-10°C) may have occurred too early to capture reproductively mature females, and its egg mass collections (May 8-23, 2000) may have occurred too late to collect many egg masses. In contrast, the GE survey was conducted between April 21 and May 8, 2003, and numerous egg masses were found in ponds with an average surface water temperature of 15°C. **Regardless of the cause of these discrepancies, the large numbers of egg masses found in 2003 indicate significant reproductive activity in the leopard frog population in the PSA.**

The ERA notes that no egg masses were found in 27 of the 44 ponds surveyed (Vol. 5, p. E-107), implying that contamination may have contributed to this finding. However, it is far more likely that habitat limitations primarily determined where leopard frogs deposited their eggs. As discussed in Attachment F, at least seven habitat factors influence the suitability of a given pond for breeding, and the initial reconnaissance for the GE survey was conducted in a conservative manner, in an effort to ensure that a broad cross-section of all ponds with potentially suitable habitat was surveyed, given the difficulty of differentiating habitat quality across these variables. In fact, a number of the ponds surveyed by GE did not have suitable habitat or environmental conditions for leopard frog breeding, based on specific factors such as water pH, temperature and permanence.

The ERA also states that the GE survey does not show how the total number of egg masses relates to overall species reproductive health (Vol. 1, p. 4-80). However, given FEL's (2002a) findings, the objective of the egg mass survey was to determine, simply and directly, whether leopard frogs are reproducing in the field under natural conditions. The results show that they are.

6.1.3 EPA's anecdotal observations

The ERA presents anecdotal information gathered during collections made for FEL leopard and wood frog studies as a separate line of evidence (Vol. 1, p. 4-65; Vol. 5, p. E-107). Specifically, the absence of leopard frog egg masses and adult females at three ponds each (as well as the absence of wood frog eggs at one vernal pool) are provided as anecdotal evidence of PCB-related impacts on frogs. No study goals or risk questions are associated with these anecdotal observations, and no additional information is provided on characteristics of these ponds that may influence their use by amphibians (e.g., pH levels, water temperature, or permanence). Moreover, a sample size of four does not offer a representative perspective of breeding activity throughout the PSA. In addition, these anecdotal observations do not acknowledge the complete absence of leopard frog adults, eggs, and larvae in the three reference ponds surveyed (Vol. 1, p. 4-65; Vol. 5, p. E-107). As a result of these issues, EPA's anecdotal observations do not warrant inclusion as a separate line of evidence for amphibians, let alone a weighting of moderate in the weight-of-evidence analysis (Vol. 1, Table 4.5-5; Vol. 5, Table E.4-5).

6.2 Site-Specific Studies on Wood Frogs

The ERA assesses risks to wood frogs (*Rana sylvatica*) using site-specific toxicity tests conducted by both EPA and GE. In addition, the ERA uses a population model to evaluate the potential population-level implications of mortality, metamorphosis, and malformations observed in the EPA study. Through selective analysis of data from the wood frog toxicity tests, the ERA overstates the risks to wood frogs resulting from exposure to PCBs.

6.2.1 EPA's wood frog study

This three-phase study, conducted in 2001 by EPA contractors FEL and Woodlot, evaluated growth, development, and maturation of wood frog larvae and metamorphs (FEL 2002b). In Phase I, egg masses were collected from 11 Housatonic River vernal pools and 3 reference pools. Egg masses and the larvae that hatched from them were exposed in the laboratory to one of three treatments: (1) water and sediment from their natal vernal pools; (2) water and sediment from the opposite types of pools (e.g., PSA pools for reference larvae and vice versa); or (3) sediment spiked with Aroclor 1260. Egg mass viability, as well as larval growth, development, and metamorphosis, were compared across treatments. In Phases II and III, larvae and metamorphs were collected from the same vernal pools at three-week intervals and were evaluated for growth, development, metamorphosis, malformations/abnormalities, and sex ratio (Phase III only).

Although this study collected data on a wide range of endpoints relevant to the survival, development, and maturation of wood frog egg masses, larvae, and metamorphs (as listed below), the only endpoints that showed significant effects of PCBs were larval malformations and metamorph abnormalities in Phases I and III and skewed sex ratios (i.e., more females than males) in Phase III. The ERA relies on these three endpoints to show adverse effects and to derive site-specific effect thresholds (see Vol. 1, p. 4-62; Vol. 5, p. E-90). The problems with the ERA's reliance on these endpoints are summarized below and discussed in more detail in Attachment G to these Comments.

Wood Frog Endpoints Evaluated in EPA Study

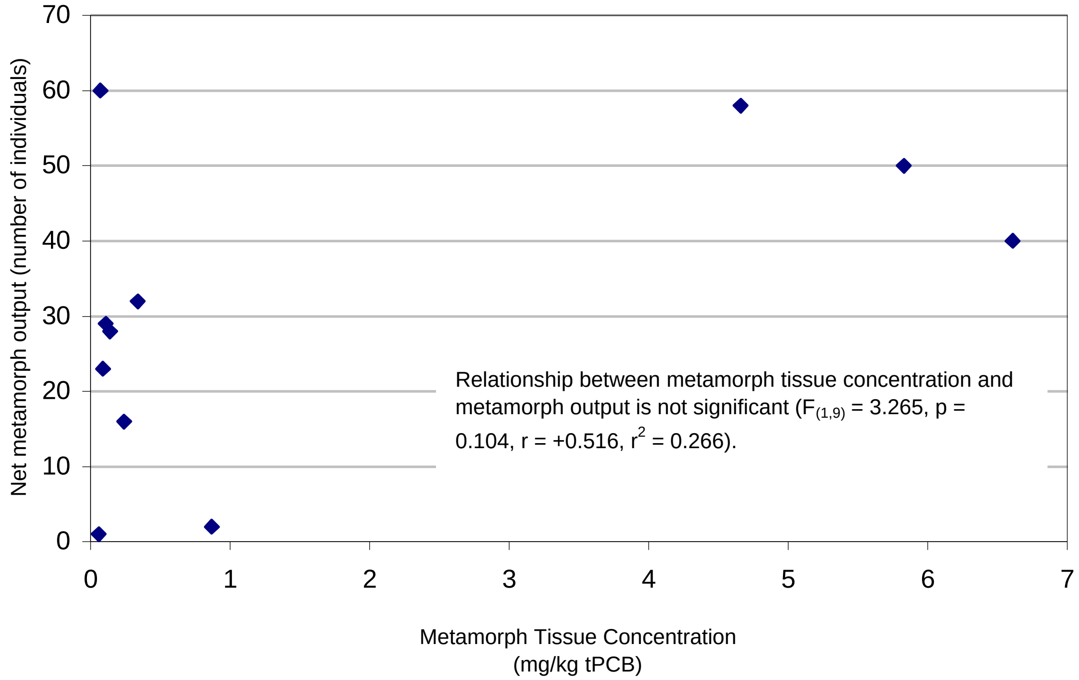
Egg Mass Weight	**Metamorphosis**
Egg Mass Count	**Metamorph Abnormalities***
Egg Mass Fertilization	**Mortality**
Egg Mass Necrosis	**Metamorph Weight**
Egg Mass Hatching Success	**Metamorph Sex Ratio***
Larval Malformations*	

Note: No significant relationships found with PCBs except for those endpoints marked with an asterisk, which indicates that a significant correlation was found for this endpoint in at least one phase.

No PCB-Related Effects on Survival, Metamorphosis, and Net Abnormality-Free Metamorph Output. In focusing on malformations/abnormalities, the ERA does not consider the data on survival, growth, or metamorphosis – endpoints clearly related to the viability of the local populations. The EPA studies showed no effects of PCBs on those endpoints (Vol. 1, pp. 4-50, 4-53; Vol. 5, p. E-80).

To better understand the combined effects of malformations, mortality, and metamorphosis on the number of normal wood frog metamorphs produced per pond, GE consultants conducted an independent regression analysis of the net abnormality-free metamorph output for target and reference ponds (i.e., the number of normal metamorphs produced). This analysis is conservative in that it assumes that all abnormalities observed will result in mortality. In fact, abnormalities may have variable implications in terms of survival, normal development, and breeding potential. For example, tail abnormalities tend to be neither fatal nor permanent (i.e., once metamorphosis is complete, the tail is reabsorbed). As a result, this independent analysis likely overestimates the potential effects of observed abnormalities. Because both the total number of larvae that completed metamorphosis and the number of metamorphs that had abnormalities are documented in the EPA Phase I studies (but not in Phase II or III), the total number of abnormality-free metamorphs produced from the eggs collected from each site could be determined for both the treatment and reference groups. GE's independent analysis, documented in Attachment G, shows that there is no significant negative relationship between net metamorph output and PCB exposure (measured as sediment, egg mass tissue concentration, or metamorph tissue concentration) in any of the Phase I studies. This is illustrated in Figure 6-1 below for the relationship between net abnormality-free metamorph output and PCB concentrations in metamorph tissue. This analysis indicates that the rates of abnormalities observed in metamorphs did not affect overall reproductive output and likely would not result in reduced recruitment of metamorphs into the population.

Figure 6-1. Regression of Net Abnormality-Free Metamorph Output vs. Tissue PCB Concentrations



Unclear Relevance of Sex Ratio Data. Review of the underlying sex ratio data indicates that those data provide no clear evidence of PCB effects. First, the results of the Phase III study are not consistent with the field data collected by Woodlot in 1999, which indicate that the sex ratios of breeding adult wood frogs were not skewed (44 to 52% female) (Vol. 5, Att. E.3, Table 5) and were within a range defined as normal by FEL (2002b p. 41, see also Gilbert et al. 1994, Merrell 1977, Reeder et al. 1998, and Stebbins and Cohen 1995). Moreover, while the ERA reports significant negative correlations between PCBs in tissue and sediment and sex ratio skewness in Phase III metamorphs (Vol. 1, Table 4.4-5), that analysis inappropriately includes a site (Site 8-VP-1) that had a sample size of only three metamorphs. Given the sample size of 3, there is a 25% likelihood that all three individuals would be of the same sex, even if the true sex ratio for the site were 1:1. GE contractors have used an ANOVA-based regression to independently analyze the sex ratio data, excluding Site 8-VP-1 (due to low sample size and because other sites with higher sample sizes had sediment PCB concentrations greater and lower than the concentration at Site 8-VP-1). This analysis, documented in Attachment G, found a statistically significant positive relationship between sediment PCB concentrations and sex ratio (i.e., percent female), but no statistically significant relationship between tissue PCB concentrations and sex ratio. Tissue concentrations represent a more direct measure of the delivered dose and, therefore, a more accurate estimate of exposure than sediment concentrations.

In any event, the ERA's analysis of sex ratio skewness neglects to acknowledge that sex ratios in amphibians can be affected by a number of environmental factors, including temperature, pH, and exposure to chemicals (Goleman et al. 2002; Holenweg and Reyer 2002; Reeder et al. 1998; Saidapur et al. 2001). Because the Phase III study evaluated metamorphs under natural environmental conditions, it is not possible to distinguish effects on sex ratios caused by natural conditions (e.g., temperature, pH) from those related to exposure to PCBs or other COPCs. Sex ratios of wood frogs are not reported for Phase I, a controlled laboratory study where exposure-response relationships could have been rigorously tested. Given these substantial issues with the sex ratio data, they do not provide reliable evidence of adverse PCB effects and should not be used to develop site-specific effects thresholds.

6.2.2 GE's wood frog study

In the spring and summer of 2001, Dr. William Resetarits of Old Dominion University and ARCADIS conducted a GE-sponsored *in situ* study of wood frogs in the PSA. A report on this study is provided in Attachment H to these Comments. This study evaluated the effects of maternal transfer of PCBs (as evidenced by concentrations in newly hatched larvae) and density-dependence (reflected by various initial densities) on larval growth, development, and metamorphosis in relatively clean ponds.[13] This study used widely accepted ecological methodologies for examination of stressor effects on the early life history of anurans to examine ecologically relevant endpoints (survival, proportion metamorphosed, mass, and time to metamorphosis) that are relevant to population-level responses. Statistical methods were sophisticated, rigorous, and powerful. The results showed strong and consistent negative effects of density on all measures of larval performance. However, the study showed no dose-response relationship between hatchling PCB levels and survival or growth, nor did it show evidence of synergistic effects of density and early-larval PCB exposure. Given the purposes of this study and its use of accepted methodologies, the ERA's weighting of the study as low (Vol. 1, Table 4.5-4) is unduly harsh.

6.2.3 EPA's population model

EPA developed a "stochastic population model" to determine whether effects of tPCBs on individual wood frogs were influencing wood frog populations in the PSA (Vol. 5, Att. E.3). According to the ERA, the model analysis shows that tPCBs hasten the decline of wood frog populations, reduce population numbers, and increase the likelihood of extinction (Vol. 1, p. 4-82; Vol. 5, p. E-123).

---

[13] This study was originally designed to assess the potential effect of sediment PCB exposure on the developing larvae as well. However, this goal could not be met because of drought conditions in the study area that prematurely reduced water levels in ponds, resulting in only two ponds, with very low levels of PCBs in sediments, retaining sufficient water to support *in situ* enclosures. As a result, the study was redesigned to evaluate only the effects of maternal transfer and density-dependence. The ERA's criticism of this study on the ground that it failed to evaluate the sediment exposure pathway (Vol. 1, p. 4-79; Vol. 5, p. E-129) does not take account of that redesigned intent.

A critique of this model is included in Attachment G to these Comments. As shown there, EPA's stochastic model is structurally flawed and does not provide realistic projections of time trends in abundance of wood frog populations. In fact, the model predicts that the entire wood frog population inhabiting the PSA has at least a 50% probability of extinction (defined as a 95% decline in the population size) within approximately 32 years in the absence of PCB exposure and 17 to 30 years with PCB exposure – and even faster (< 6 years) if the population is assumed to be already declining (Vol. 5, Att. E-3, Table 22). Since PCBs have been present in the PSA since the 1930s, it follows that if EPA's model were accurate, there should be no wood frogs present in the PSA, which is demonstrably not the case.

The principal reason for the unrealistic model projections is that, contrary to the methods described in the text (Vol. 5, Att. E.3, p. 17), EPA's model does not consider density-dependent regulatory processes that promote the rapid growth of amphibian populations at low population sizes. Examples of such processes include density-dependent egg production by adults and density-dependent growth, survival, and metamorphosis of tadpoles. The operation of these processes has been demonstrated in many amphibian populations, including the study used by EPA as a source of parameter values for its stochastic population model (Berven 1990). In addition, GE's wood frog study (see Section 6.2.2) demonstrates the operation of density-dependent processes in wood frog populations inhabiting vernal ponds within the PSA. Because it does not account for these processes, EPA's model greatly overstates risks of extinction for both exposed and unexposed wood frog populations. An additional problem with this model is that the exposure values used for the mortality and metamorphosis regressions are incorrect, as also shown in Attachment G.

6.3 EPA's Vernal Pool Study

EPA contractors (Woodlot Alternatives) conducted an amphibian community study with the objective of documenting the number of frogs entering and the number of frogs and metamorphs leaving four vernal pools. It also documented species abundance, richness, and malformation rates in those four vernal pools. Species density and richness were assessed relative to PCB sediment concentrations in the vernal pools.

No adverse effects were observed for adult deformities, timing of breeding, body sizes, courtship and breeding behavior and condition, 10-day survival of larvae, larval growth rates, larval malformation rate or metamorphosis (Woodlot 2003). Species density and richness were reported to be lower in ponds with higher sediment tPCB concentrations (Vol. 1, p. p. 4-64; Vol. 5, p. E-106). However, the study is limited by a small sample size and fails to consider the potential influence of confounding factors that might influence these measures. For example, several key pool characteristics, such as pond depth, pH, temperature, and presence/absence of fish, are not reported. However, pool surface area is reported (Woodlot 2003, Table 3), demonstrating an inverse relationship between the size of the pool and the concentration of PCBs in sediment. This area effect might well have explained more of the variability in density and richness observed across the ponds than could be explained by PCBs in sediment.

In addition, malformation rates in larval wood frogs were reported to be high in all PSA ponds irrespective of sediment tPCB concentrations, which range from 0.72 to 32.3 mg/kg (Vol. 1, p. 4-64; Vol. 5, p. E-106). However, no exposure-response relationship was observed for malformations across this large concentration range. Moreover, because no data were collected from reference ponds, it is not possible to determine if the malformations reflected a regional trend. In any case, the low rates of malformations in metamorph and adult wood frogs (Vol. 1, p. 4-64; Vol. 5, p. E-106) indicate that these malformations do not translate to later developmental stages. Thus, any conclusions regarding the effects of tPCBs on the parameters measured in this study should be considered extremely uncertain.

6.4 Unrealistically Conservative Effects Thresholds

The ERA presents a site-specific MATC for sediment/soil (3 mg/kg tPCBs) based on Phase III malformation and sex ratio results (Vol. 1, p. 4-63; Vol. 5, pp. E-104 - E-105, Table E.4-1). As discussed above (Section 6.2.1), the skewed sex ratio endpoint does not provide a reasonable basis for establishing an effects threshold, because there is no evidence that adult wood frogs have skewed sex ratios and because there was no significant relationship between tissue PCB concentrations and skewed sex ratios in the Phase III metamorphs. Moreover, the sediment MATC of 3 mg/kg tPCB is slightly less than the EC20 for Phase III malformations (3.27 mg/kg tPCB, spatially weighted) and more than an order of magnitude lower than the EC20 and EC50 for Phase I malformations (both of which are >32.3 mg/kg tPCB, spatially weighted, and 62 mg/kg tPCB based on vernal pool/synoptic data) and the EC50 for Phase III malformations (38.6 mg/kg tPCB, spatially weighted) (Vol. 5, Table E-4.1).[14] The EC50 for Phase III malformations (38.6 mg/kg tPCBs) would provide a more appropriate MATC. However, even that threshold likely overestimates population risks because: (1) it is based on the Phase III analysis that included both internal and external malformations, and (2) there is no evidence that the malformation rates affect survival, metamorphosis or net abnormality-free metamorph output.

[14] The vernal pool/synoptic sediment tPCB concentrations should be used for the Phase I study because they represent the sediment tPCB concentrations to which the eggs and larvae were exposed in the laboratory through metamorphosis. Spatially weighted concentrations are more relevant for Phase III because metamorphs were exposed only to sediment in their native pools and not in the laboratory.

The ERA also specifies a site-specific MATC of 1 mg/kg tPCBs in frog tissue based on the data from its wood frog toxicity study. That value appears to be based on the EC20 for sex ratio (Vol. 5, Table E.4-1), which is not a reasonable endpoint for an effects threshold, as previously discussed (Section 6.2.1). It is approximately six times lower than the EC20 for Phase I malformations (>6.61 mg/kg) and the EC50 for Phase III malformations (5.65 to 6.54 mg/kg) (Vol. 5, Table E.4-1). While the latter may be considered a more appropriate tissue MATC, it would likely overestimate population risks to frogs for the same reasons given in the previous paragraph for the sediment/soil MATC based on the same endpoint (i.e., the EC50 for Phase III malformations).

The ERA also supports the tissue MATC of 1 mg/kg tPCBs based on a review of the literature (Vol. 1, pp. 4-58 - 4-60; Vol. 5, p. E-89, p. E-108). However, the two tPCB LOAELs reported as being below 1 mg/kg are not supported by the underlying study (Gutleb et al. 2000). One of these effects thresholds, developed for the time to metamorphosis for the African clawed frog, is based on exposure to the most potent PCB congener (PCB 126) rather than tPCBs. Hence, this LOAEL is neither species-specific nor stressor-specific and should not be considered in the development of the MATC. The second LOAEL of 0.24 mg/kg tPCBs (Gutleb et al. 2000) is based on body weights in exposed European common frogs that are greater, not less, than controls. This effect is neither consistent with PCB-induced toxicity nor adverse. For these reasons, this LOAEL also should not be considered in the development of the literature-based effects threshold. When these two LOAELs are removed, the lowest LOAEL is 6 mg/kg tPCBs (Savage et al. 2002) (see ERA, Vol. 1, Fig. 4.4-11), which is consistent with the tissue MATC derived from the EC50 for Phase III malformations in the site-specific study. However, there is no point in using such a literature-based value when site-specific data are available.

6.5 Overall Assessment

The ERA concludes that there is a significant risk to frog species in the PSA. Confidence in this conclusion is given as moderate to high, based on the asserted concordance in predictions of risk from multiple measurement endpoints (Vol. 1, p. 4-82; Vol. 5, p. E-121). These conclusions are primarily based on the results of the two EPA toxicity studies. As shown above, however, the ERA overstates the conclusions from these two studies. Due to the serious limitations of the leopard frog study, that study does not provide reliable evidence of harm. Moreover, the ERA's interpretation of EPA's wood frog study overemphasizes endpoints (i.e. malformations and sex ratio) that are not directly relevant to the sustainability of the local population, while failing to consider more ecologically relevant endpoints such as survival, metamorphosis, and reproductive output. In addition, the ERA's conclusions do not adequately consider evidence from EPA's and GE's field studies that documented reproducing amphibian populations in the PSA, and it uses a flawed and unrealistic model to extrapolate individual-level effects to effects on local populations. Further, the site-specific tissue and sediment effects thresholds are extremely conservative and are based on only a small subset of the available effects data. Similarly, the literature-based tissue effects threshold is based on a misinterpretation of the underlying studies, resulting in an effects threshold six times lower than the lowest value supported by the underlying data. Consequently, the HQs that depend upon these thresholds overstate risk.

The weight-of-evidence evaluation for amphibians is biased in that it does not present a balanced assessment of EPA's and GE's studies. For example, EPA's anecdotal observations are accorded Moderate weight, whereas GE's wood frog study is given Low weight and GE's leopard frog egg mass study is given Low weight (Vol. 1, p. 4-74; Vol. 5, Table E.4-5). While there are limitations associated with GE's studies, specific criticisms leveled against these studies in the ERA reflect misunderstandings of the underlying studies and/or biased treatment of study weaknesses for individual attributes. By contrast, EPA's anecdotal observations have no associated risk-related questions, statistical analyses, or quality control measures, and relate only to alleged effects in the PSA, failing to address any effects in the reference areas. Similarly, despite the substantial problems with the EPA leopard frog study, it is given Moderate/High weight (the same weight given to the EPA wood frog study).

Based on the considerations described above, it can be concluded that: (a) while there is evidence of increased larval and metamorph abnormalities among amphibians in the PSA, those levels of abnormalities do not appear sufficient to affect the net output of metamorphs without abnormalities, and thus the populations; and (b) the field data provide evidence of reproducing populations of amphibians in the PSA.

SUMMARY OF CONCLUSIONS
RISKS TO AMPHIBIANS IN PSA

While EPA's wood frog toxicity study provides some evidence of increased larval and metamorph abnormalities among amphibians in the PSA, the site-specific toxicity and field data do not indicate that these abnormalities translate into adverse effects on local amphibian populations.

6.6 Extrapolation to Downstream Reaches

The ERA assesses risk in reaches downstream of the PSA based on the sediment/soil MATC of 3 mg/kg tPCBs. Problems with this MATC that result in an overestimation of risk are described in Section 6.3 above. Moreover, this MATC is applied to floodplain soils and sediment downstream of the PSA, regardless of habitat. The majority of the habitat below the PSA is, in fact, not suitable for amphibian populations, as there are few wetland areas available. Because the site-specific effects thresholds used to derive the MATC are based on effects in larvae (which require standing water), the MATC should not be applied to areas of the floodplain impacted by PCBs where wetlands and vernal pools are not present.

Maps of wetlands and vernal pools downstream of the PSA are available through MassGIS (www.state.ma.us/mgis). These maps have been downloaded and are shown on Figures 6-2a through 6-2c. Based on review of these maps, the 100-year floodplain between Woods Pond and Rising Pond contains approximately 184 acres of wetlands, representing approximately 8% of the 2,448 acres in this area. South of Rising Pond to the Connecticut border, the 100-year floodplain contains approximately 193 acres of wetlands, representing approximately 4% of the total 5,028 acres. Because of both the conservative MATC and the limited amphibian habitat present downstream from Woods Pond, the ERA clearly overestimates the extent of downstream risks to amphibian populations and communities.

See Table 6-1 for specific comments on Charge Questions 3.2(a)-(j)

Table 6-1. Assessment Endpoint: Community Condition, Survival, Reproduction, Development, and Maturation of Amphibians

EPA Charge Question #3.2	GE Response
(a) Were the EPA studies and analyses performed (e.g., field study, site-specific toxicity study, comparison of exposure and effects) appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), and based on accepted scientific practices?	**Leopard Frog Study:** The use of frogs purchased from a biological supply company as "references" is inappropriate and renders the study's conclusions based on comparisons to those frogs unreliable. (Section 6.1.1) **Wood Frog Study:** Study design was appropriate. **Vernal Pool Study:** Basic methodology was appropriate, but the sample size (n=4) was low. (Section 6.3) **Anecdotal Observations:** It is inappropriate for anecdotal observations to be presented as a separate line of evidence. (Section 6.1.3) **Comparisons to Effects Thresholds (HQs):** Site-specific effects thresholds are based on inappropriate endpoints. Literature-based effects threshold is unnecessary given site data, and misinterprets the literature. (See below.)
(b) Were the GE studies and analyses performed outside of the framework of the ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), based on accepted scientific practices, and incorporated appropriately in the ERA?	**GE Leopard Frog Egg Mass Survey:** This survey was appropriate for its purpose – documenting existence of leopard frog reproductive activity in the PSA. The results of this study are misinterpreted and given a lower weight than appropriate in the ERA. (Sections 6.1.2) **GE Wood Frog Study:** While this study did not consider effects of sediment PCB exposure, it did evaluate effects of PCB exposure to larvae via maternal transfer. Given the purpose of this study and its use of accepted methodologies, the ERA's weighting of the study is too low. (Section 6.2.2)
(c) Were the estimates of exposure appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?	**EPA Leopard Frog Study:** No comments on exposure estimates. **EPA Wood Frog Study:** Use of spatially weighted averages instead of synoptic sediment data is appropriate for Phase II and III; however, synoptic sediment data should be used for Phase I because organisms were exposed only to "synoptic" sediment in the laboratory. (Section 6.4, note 14) **EPA Vernal Pool Study:** No comments on exposure estimates.
(d) Were the effects metrics that were identified and used objective, reasonable, transparent, and appropriate?	**Site-Specific Effects Metrics:** The site-specific effects metrics of 3 mg/kg tPCB in sediment and 1 mg/kg in tissue are based on the sex ratio data from EPA's wood frog study, which do not provide a reasonable basis for establishing effects thresholds. The EPA malformation data would support tPCB effects thresholds of 38 mg/kg in sediment and 6 mg/kg in tissue, but these thresholds likely overstate population risks. (Section 6.4) **Literature-Based Effects Metric:** The literature-based effects threshold of 1 mg/kg tPCBs in tissue is based on an inappropriate interpretation of the underlying studies and is considerably lower than supported by the data. (Section 6.4)
(e) Were the statistical techniques used clearly described, appropriate (i.e., objective, consistent, and reasonable), and properly applied for the objectives of the analysis?	**EPA Leopard Frog Study:** The ERA's comparisons to commercially purchased frogs are inappropriate. (Section 6.1.1) **EPA Wood Frog Study:** Individual probit analyses used to derive the EC50 were not presented and therefore could not be validated. Otherwise, the methods appear to be appropriate. **EPA Vernal Pool Study:** Statistical analyses failed to account for likely area effect that resulted from the inverse relationship between pond size and sediment concentration. (Section 6.3)

Table 6-1. Assessment Endpoint: Community Condition, Survival, Reproduction, Development, and Maturation of Amphibians

EPA Charge Question #3.2	GE Response
(f) Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable)?	**Leopard Frogs:** The ERA overstates risks to leopard frogs in the PSA because it places too much weight on EPA's leopard frog study, which does not provide reliable evidence of impaired reproduction or development (Section 6.1.1), and does not adequately take account of GE's leopard frog egg mass survey, which shows significant reproductive activity in the leopard frog population in the PSA (Section 6.1.2). **Wood Frogs:** The ERA overstates risks to wood frogs in the PSA because it overemphasizes endpoints (i.e., malformations and sex ratio) that are not directly relevant to the sustainability of the local population, while failing to consider more ecologically relevant endpoints such as survival, metamorphosis, and reproductive output, which showed no adverse PCB effects (Section 6.2.1). In addition, the ERA uses a flawed and unrealistic model to predict population-level effects (Section 6.2.3). **Vernal Pool Study:** The ERA overstates the implications of this study by failing to take account of the low sample size, potentially confounding habitat effects, and lack of exposure-response for metamorph malformations. (Section 6.3) **HQs:** The ERA's HQs overestimate risks by at least 10 times for sediment and 6 times for tissue due to inappropriate effects metrics. (Section 6.4)
(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements could be made.	**EPA Leopard Frog Study:** Uncertainties associated with the use of commercially purchased frogs as a basis of comparison with PSA frogs are not sufficiently addressed. (Section 6.1.1) **EPA Wood Frog Study:** Uncertainties associated with extrapolating from malformation data to potential impacts on wood frog populations, despite the lack of effects on metamorphosis and survival, are not adequately addressed. Uncertainties in sex ratio data are not adequately addressed. (Section 6.2.1) **EPA Vernal Pool Study:** Uncertainties associated with low sample size and potential area effect on community measures are not adequately recognized. (Section 6.3)
(h) Was the weight of evidence analysis appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable)? If not, how could it be improved?	**EPA Leopard Frog study:** This study is given too much weight (Moderate/High) given the considerable problems with this study. (Sections 6.1.1 and 6.5) **GE Leopard Frog Egg Mass Survey:** The ERA's criticisms of this study reflect misunderstandings of the study. Weight given to this study (Low/Moderate) is too low. (Sections 6.1.2 and 6.5) **EPA Anecdotal Observations:** Weight given to these observations (Moderate) is far too high. (Sections 6.1.3 and 6.5) **EPA Wood Frog Study:** Weight given to this study (Moderate/High) is appropriate, but the data are over-interpreted. **EPA Vernal Pool Study:** Weight given to this study (Moderate/High) is too high given limitations. (Section 6.3)
(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?	The ERA overestimates risks to downstream reaches by relying on an overly conservative MATC and by applying that MATC to the entire 100-year floodplain when only a small fraction of that floodplain contains suitable habitat for amphibian larvae (the life stage on which the MATC is based). (Section 6.6)
(j) In the Panel members' opinion, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?	**The ERA overstates risks to local populations of frogs.** While EPA's wood frog study provides some evidence of increased larval and metamorph abnormalities among frogs in the PSA, both the toxicity test data and the site-specific field data indicate that these abnormalities do not adversely affect the local frog populations. Contrary to EPA's flawed population model, the field data show reproducing populations of frogs in the PSA. (Section 6.5)

7. FISH (Question 3.3)

Key Points

- Field studies conducted by both EPA and GE showed that there are self-sustaining populations and communities of fish in the Housatonic River, including the PSA. GE's study further showed that the largemouth bass population in the river is reproducing and has population parameters within the range that would be expected in a similar system in the Northeast. The ERA's criticisms of that study are unwarranted.

- The ERA misinterprets the results of EPA's site-specific Phase I and II toxicity studies by concluding that they showed PCB-related toxicosis. A careful review of the data from those studies reveals that they did not show consistent relationships between PCB exposure and adverse effects.

    Ø While these studies did find various statistically significant differences between at least one Housatonic River site and the reference site for a number of adult and offspring endpoints, those differences were not consistent either spatially (among sites) or temporally (between developmental stages) or between the Phase I and II studies, and did not show clear exposure-response relationships with PCBs.

- The tissue-based effects metrics derived from these studies are overly conservative for the same reasons noted above and also because the effects metric derived from the Phase II studies: (a) was based on pooled endpoints that individually did not show consistent PCB dose-response relationships; (b) combined results for coldwater species (salmonids) and warmwater species, when the former are known to be more sensitive to PCBs; and (c) used a conversion factor for converting from egg concentrations to whole body tissue concentrations that is unsupported by the overall data.

- The literature-based tissue effects thresholds are likewise overly conservative because they: (a) inappropriately included studies of a species (lake trout) that is not found in the Housatonic River and is more sensitive to PCBs than the species that do occur in the river; (b) are based (for tPCBs) on an inappropriate averaging of different effects metrics in different tissue types; and (c) utilized the egg to whole body conversion factor that is unsupported by the literature.

- The ERA's weight-of-evidence evaluation incorrectly asserts that the GE and EPA field studies showed undetermined evidence of harm, when they in fact showed no evidence of harm. It also does not adequately take account of the uncertainties associated with the comparisons to literature-based thresholds and thus assigns them too high weight.

- **The site-specific toxicity and field studies do not support the ERA's conclusion that there are "ecologically significant" risks to fish from PCBs and TEQs in the PSA in terms of population sustainability. The data from those studies, considered together, indicate that, while some effects may be occurring sporadically in individual fish, such effects are not adversely affecting the local fish populations and communities. Thus, there are negligible risks to the populations and communities of fish in the PSA.**

7. FISH (Question 3.3)

The ERA employs three types of measurement endpoints in its evaluation of potential risks to fish: (1) a two-phase site-specific toxicity study conducted by EPA contractors; (2) field surveys conducted by both EPA and GE contractors; and (3) comparisons of measured tPCB and TEQ concentrations in fish tissue to literature-based and site-specific effects metrics. Based on these endpoints, the ERA concludes, with moderate confidence, that there are "ecologically significant" but low-magnitude risks to fish in the PSA from both tPCBs and PCB-related TEQs, and that there are marginal and uncertain risks to coldwater fish (but not warmwater fish) downstream of the PSA (Vol. 1, pp. 5-74 - 5-75). GE believes that these conclusions overstate the risks to fish in the Housatonic River. Specific concerns related to the interpretation of these studies are discussed below and in Attachment I. A summary of specific comments related to Questions 3.3(a)-(j) are provided in Table 7-1 (at the end of this section).

7.1 EPA's Site-Specific Toxicity Studies

EPA contractors conducted a two-phase study to evaluate reproductive toxicity of site-specific COPCs to fish (Tillit et al. 2003a,b).[15] The Phase I study quantified PCB concentrations in Housatonic River adult largemouth bass and evaluated effects in their offspring (Tillitt et al. 2003a). In this Phase I study, adult largemouth bass were collected from three locations in the Housatonic River and from a reference location (Threemile Pond). They were transported to a laboratory, where a subset was sacrificed to determine tissue chemistry and the remainder was raised in experimental ponds through spawning. The adult bass were evaluated both pre-spawn and post-spawn for biochemical and hormonal endpoints, organ and gonadal histology, and body and organ sizes and weights. Their offspring were monitored for survival, growth, developmental parameters, and cytochrome P40 induction.

The Phase II studies were designed to test whether PCBs are causally linked to the endpoints evaluated in Phase I (Tillitt et al. 2003b). Brood fish from the Phase I study were sacrificed to obtain chemical extracts for subsequent injection into eggs in Phase II. The extracts were characterized with respect to tPCB and TEQ concentrations to allow evaluation of dose-response relationships. Eggs from non-native largemouth bass, medaka, and rainbow trout were injected with a range of concentrations of these extracts, as well as with standards – 2,3,7,8-TCDD and 3,3',4,4',5-pentachlorobiphenyl (PCB 126). These standards were chosen because they are known to cause aryl hydrocarbon receptor (Ah-R)-mediated toxicity (i.e., dioxin-like toxicity). Additional negative control treatments included uninjected

[15] Section 5 (Volume I) of the ERA refers to Tillett et al. 2001 and 2002, which were draft versions of these reports. Appendix E (Volume 5) appropriately refers to Tillitt et al. 2003a and 2003b. For the purpose of these Comments, we assume that the ERA has relied upon the final reports (i.e., Tillitt et al. 2003a and 2003b).

eggs and eggs injected with the solvent that was used as a carrier for the extracts (i.e., sham-injected). The treated eggs and the fry that hatched from them were then reared in the laboratory and monitored for the same endpoints evaluated in Phase I.

The ERA concludes that the Phase I and II studies provide evidence of PCB-related toxicosis (Vol. 1, pp. 5-30, 5-38; Vol. 5, pp. F-32, F-41). Effects thresholds are derived based on the results of both studies (Vol. 1, pp. 5-39 - 5-45; Vol. 5, pp. F-42 - F-47). However, GE has found that the incidence of statistically significant effects is inconsistent both within and between the two phases, and that a number of the effects that were observed are not unique to PCB toxicosis. A clear demonstration of PCB-related toxicity would require: (1) that the observed effects are consistently and statistically greater in Housatonic River exposed fish than in reference fish from Threemile Pond or in the uninjected and sham-injected controls (Phase II only); (2) a clear and repeatable exposure- or dose-response relationship (i.e., higher effects in sites with higher exposures to tPCBs or TEQs); and (3) consistent effects across the Phase I and II studies. The results of the Phase I and Phase II studies do not meet these criteria and do not provide evidence of PCB-related toxicosis. These issues are summarized below and a detailed discussion is provided in Attachment I (Sections 3 through 5).

7.1.1 Lack of consistent evidence of dose-dependent PCB effects

Phase I

Although Phase I found statistically significant differences between at least one of the Housatonic River sites and the reference site for a number of adult and offspring endpoints (Vol. 1, pp. 5-29, 5-30; Vol. 5, pp. F-30 - F-32), the incidence and magnitude of these various effects are not consistent either spatially (i.e., among Housatonic River sites) or temporally (i.e., between developmental stages). Significant differences from the reference site were not always found in the sites with the highest tPCB or TEQ concentrations and the overall patterns of observed effects do not indicate an exposure-response relationship with PCBs. While some of the effects observed may be considered part of the “suite” of effects associated with PCB-toxicosis (e.g., swim bladder and craniofacial deformities), a number of endpoints evaluated are general indicators of stress and are not specifically associated with PCB-related toxicosis (see Attachment I, Section 3). The significant incidence of abnormal swim bladders in fish from Housatonic River sites occurred in only a subset of spawns and did not have an effect on survival (see Section 7.1.2 below).

Phase II

In the Phase II study, the majority of statistically significant lethal and sublethal effects were observed in the treatments that used the PCB 126 and 2,3,7,8-TCDD standards (Vol. 1, pp. 5-32 - 5-38; Vol. 5, p. F-

36 - F-41). In many cases, these same effects were not observed in fish injected with Housatonic River extracts or in the Phase I study. While a subset of endpoints, including most abnormalities, had a higher occurrence of effects in at least one Housatonic River treatment than in the reference or controls, and, in some cases, individual trials displayed dose-related effects, these effects were generally not consistent within or across Housatonic River trials (see Attachment I, Section 4).

In order to find effects, the ERA pools all the individual abnormality data to assess an endpoint of "total abnormalities" (Vol. 1, p. 5-34; Vol. 5, p. F-39). In addition, to calculate ED50s for each species, the ERA combines all lethal and sublethal effects (i.e., mortality or one or more abnormalities) (Vol. 1, pp. 5-40, 5-41; Vol. 5, p. F-43). As discussed in Attachment I (Section 4.2), this pooling of effects data is inappropriate because it results in a determination of significant dose-related effects compared to controls when no or weak dose-response relationships are demonstrated for the individual abnormalities. Therefore, the effects thresholds (ED50s) for the Phase II study do not represent PCB-related effects and should not be used to assess risk.

This is further demonstrated by the fact that when the effects levels (LD50s and ED50s) derived from the Phase II study are normalized to TEQs, extracts from Housatonic River fish appear more toxic than the 2,3,7,8-TCDD and PCB 126 standards. Such a result is inconsistent with studies that have evaluated the relative toxicities of chemicals acting via the Ah-R, which show that 2,3,7,8-TCDD is the most potent of the dioxin-like chemicals and that PCB 126 is more toxic than the typical PCB mixtures such as those found in the Housatonic River (Van den Berg et al. 1998; Wright and Tillitt 1999). There is no support for the ERA's speculation that this finding is due to the synergy of PCB mixtures or due to other PCB compounds not quantified in the TEQ model (Vol. 1, p. 5-38; Vol. 5, p. F-41), rather than the result of the ERA's flawed method of calculating the effects levels.

7.1.2 No dose-response relationship between PCBs and swim bladder deformities

The ERA reports that, in the Phase I study, there were statistically significant increases in swim bladder deformities at all Housatonic River sites relative to the reference site at 15-days post swim-up (Vol. 5, p. F-32). However, most of these abnormalities occurred in only one spawn at each location (Tillitt et al. 2003b). The relevance of these abnormalities is unclear, as there is no apparent relationship between the incidence of swim bladder abnormalities and survival. Specifically, although Deep Reach and Woods Pond fish had significantly more swim bladder abnormalities than the reference site fish (Vol. 1, p. 5-30; Vol. 5, p. F-32), fish from neither of those Housatonic River sites showed a statistically significant reduction in survival at 15-days post swim-up compared to the reference fish (Tillitt et al. 2003a, Table 15). Further, in contrast to the results presented for Phase I, the Phase II results at the same stage (i.e., 15 days post swim-up) showed few swim bladder deformities. At swim-up, however, when most swim bladder deformities were observed in Phase II, they occurred in all treatments (including the reference) and there was no evidence of a dose-related response (Tillitt et al. 2003b, Tables 22 and 23).

This inconsistent occurrence of swim bladder abnormalities provides no evidence of a consistently higher incidence of effects in the Housatonic River sites relative to the reference site or the controls and provides no evidence of a dose-related response. Moreover, this effect was not consistently seen in the standards, indicating that it may not be associated with Ah-R-mediated toxicity and is likely associated with a factor unrelated to PCBs. For instance, failure of swim bladder inflation is a common problem in the culture of striped bass, as well as several other fish species (Chapman 1992; Colesante et al. 1986; Tucker 1987), often occurring in the majority of larvae.

7.2 Site-Specific Fish Community and Population Studies

Both EPA and GE contractors conducted field surveys of the fish community in the PSA. In addition, the GE study evaluated reproduction and habitat usage of the largemouth bass. These studies used similar methods (e.g., electrofishing, weight and length measurements, collection of scales and otoliths) to develop estimates of condition, age, species abundance, and community composition. Both studies concluded that the representative species evaluated were found in suitable habitat and are self-sustaining (ERA, Vol. 1, p. 5-51; Vol. 5, p. F-56) and that there were no detectable impacts on the fish community. The GE study also concluded that largemouth bass populations are stable and reproducing and that key population and community parameters are within the range that would be expected in a Northeast system similar to the Housatonic River. The full report on the GE study (R2 Resource Consultants 2002) is included on the compact disk provided with the ERA that contains GE's studies; a more abbreviated manuscript on the study is provided in Attachment J to these Comments. Specifically, this study found:

- There was considerable largemouth bass reproductive activity in the mainstem of the Housatonic River in the PSA, as evidenced by the number and density of active nests observed (71 separate active nests) and the observations of broods in all sites monitored.

- Estimates of young-of-year (YOY) abundance based on electrofishing-derived catch per unit effort (CPUE) varied from year to year and were comparable to estimates from other studies.

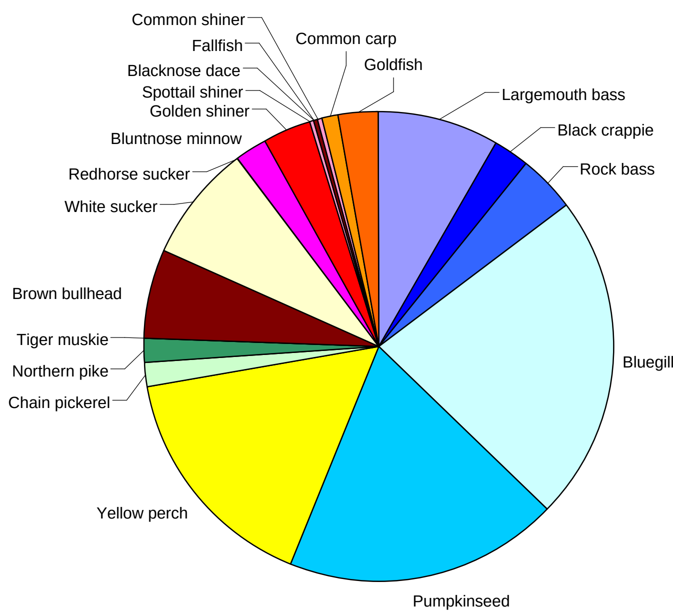
- The growth rate of YOY bass in the Housatonic River (0.92 mm/day over the first 8 weeks) was consistent with, and in some cases on the high end of, growth rates reported for other populations.

- The overall growth rate from Housatonic River largemouth bass ranked in the 90th percentile of all populations evaluated (estimated by the Brody growth coefficient).

- Mean relative weights (a condition index) computed for Housatonic River largemouth bass were above the "standard" range and within the upper 80th to 95th percentile of what could be expected from all populations evaluated.

- The overall fish community in the mainstem of the Housatonic River in the PSA (illustrated on Figure 7-1 below) includes the presence of numerous species and is more representative of a Massachusetts lake than a river, which is expected given that that section of the river is relatively flat and contains numerous backwaters and that much of that section has an "impounded" nature due to the influence of Woods Pond Dam.

Figure 7-1. Fish Community Structure in the Mainstem of the Housatonic River



The ERA misinterprets some key findings of the GE largemouth bass study. For instance, it states that the nesting survey did not show strong reproductive success (Vol. 1, p. 5-52; Vol. 5, p. F-59). In fact, the study did not assess this parameter because individual nests were not followed over time. Moreover, it is inappropriate to assume that the nests where no eggs or fry were observed were not productive. The report (Attachment J) states only that they were not observed and provides a number of possible reasons for this (e.g., already moved off the nest). The ERA also notes that this study showed lower growth of the YOY bass than in other systems (Vol. 5, p. F-59). While the total growth (length) at the end of the first growing season was lower than in many other systems, the overall growth rates of the fish were comparable to other systems. The smaller total length in the fall likely reflects a shorter growing season than most other largemouth bass systems (see Attachment J). Additionally, the ERA's conclusion that the "apparent self-sustaining nature of this population may be more a function of the low mortality rate of the adults rather than high reproductive output" (Vol. 1, p. 5-53; Vol. 5, p. F-59) is unsubstantiated. In fact, the GE study used observations of local reproductive success (qualitative observations) and end-of-growing-season YOY catch per unit effort as indicators of successful recruitment to the population. Adult mortality rates were not considered in this analysis.

The ERA also states that GE's largemouth bass study did not address the extent of tributary recruitment or the effects of PCBs on population strength or viability (Vol. 1, p. 5-52; Vol. 5, p. F-61). It speculates that if there were fishing pressure or other environmental stress in the area, the population-level impacts of the developmental effects observed in the Phase I and II toxicity studies "may be more severe" (Vol. 5, p. F-61). In fact, in addition to showing reproduction in the mainstem of the river, the study also showed that there is very limited or unsuitable largemouth bass habitat in the tributaries (R2 Resource Consultants 2002). Hence, the only way that the PSA portion of the river could support a largemouth bass population is through internal reproduction. Accordingly, any individual-level effects that may be occurring in largemouth bass in the PSA do not appear to be affecting the sustainability of the local bass population. This is not surprising given the resiliency of fish populations typically controlled by density-dependent factors. Largemouth bass produce prolific numbers of progeny that are naturally subjected to high rates of mortality throughout their early life stage (Winemiller and Rose 1992). For these populations, individual-level effects of PCBs may not result in detectable adverse effects on the population.

In short, given that PCBs have been present in the PSA for many decades, coupled with the lack of adequate habitat for largemouth bass in the tributaries, GE's study finding a healthy and reproducing largemouth bass population in the PSA provides strong evidence that PCBs are not adversely affecting that local population.

Thus, this study, as well as EPA's fish community study, found no evidence of harm to the fish population and community in the PSA. Accordingly, the ERA should characterize the evidence of harm from these studies as none rather than undetermined (Vol. 1, p. 5-70; Vol. 5, Tables F.4-6, F.4-7).

7.3 Tissue-Based Effects Thresholds

The ERA derives effects thresholds for fish both based on its two-phase toxicity studies and based on review of the literature. These thresholds are overly conservative due to the methods used to calculate them and a flawed interpretation of the underlying studies. These points are summarized briefly below and demonstrated further in Attachment I.

7.3.1 Effects metrics based on site-specific toxicity studies

The ERA derives a tissue-based effects metric of 49 mg/kg tPCBs for use in its HQ analysis and as a MATC for warmwater fish (Vol. 1, p. 5-73). This value is based on an average of its calculated tissue-based effects metrics for Phase I and Phase II, which are, respectively, <45 mg/kg tPCBs (Vol. 1, p. 5-39) and 52 mg/kg tPCBs (Vol. 1, p. 5-45). As shown in Attachment I (Section 6), these thresholds are too low for the following reasons:

- The threshold for Phase I is reported as an unbounded LOAEL for reduced survival and increased abnormalities (Vol. 1, p. 5-39; Vol. 5, p. F-42). This LOAEL is not warranted because (as discussed above) these effects did not occur consistently in Housatonic River sites, did not have a consistent exposure-response relationship with PCBs, and were not found to be PCB-related in the Phase II study.

- For Phase II, the effects level is based on the mean ED50 value for tPCB toxicity to eggs in Phase II, with application of a factor of 0.5 to convert from egg concentrations to whole body concentrations (Vol. 1, p. 5-45; Vol. 5, pp. F-47 - F-50). This effects level is not appropriate because: (1) the egg-based threshold was based on pooled endpoints (mortality and abnormalities) that individually did not demonstrate consistent dose-response relationships with PCBs, and is inconsistent with current scientific understanding of the relative toxicity of PCB mixtures compared to 2,3,7,8-TCDD; (2) that threshold was based on combined results for coldwater species (salmonids) and warmwater species, when the former are known to be more sensitive to PCBs; and (3) the use of a 0.5 conversion factor to convert from egg concentrations to whole body concentrations is not supported by the available data, which indicate that no such conversion factor is necessary.

The ERA also develops a tissue-based MATC of 12 mg/kg tPCBs for coldwater fish based on dividing the warmwater fish MATC by a factor of four (Vol. 1, p. 5-74; Vol. 5, p. F-77). There is no quantitative basis for selection of that factor. The MATC for coldwater fish should be based on the data for such fish, without application of an arbitrary factor.

7.3.2 Effects thresholds derived from literature

The ERA also compares fish tissue chemistry data to literature-based effects metrics for both PCBs and TEQs as a separate line of evidence (Vol. 1, pp. 5-22 - 5-28; Vol. 5, pp. F-25 - F-30). The effects metrics selected are 31 mg/kg for tPCBs and 50 ng/kg for TEQs, both of which are calculated by deriving thresholds for eggs and/or sac-fry/juveniles and then applying a conversion factor of 0.5. As shown in Attachment I (Section 7), these literature-based effects metrics are overly conservative because:

- They include studies of salmonid species, which have been shown to be highly sensitive to PCBs, including lake trout, which is not found in the Housatonic River and is more sensitive to PCBs than the species that do occur in the river;

- The threshold for tPCBs was based on an inappropriate averaging of different effects metrics in different tissue types; and

- Both thresholds were based on use of an egg/sac fry to whole body conversion factor that, based on the literature, is unnecessary.

7.4 Overall Assessment

The ERA concludes, with moderate confidence, that risks posed by tPCBs and TEQs to fish in the Housatonic River are "ecologically significant" (a term which it does not define) but low in magnitude (Vol. 1, p. 5-74). It reports that strength in the conclusions is derived from the concordance in predictions of risk from multiple measurement endpoints, although some uncertainty is associated with several of the endpoints (Vol. 1, p. 5-75).

GE does not agree that the available data show evidence of "ecologically significant" risks to fish in the PSA, assuming that ecological significance refers to effects on population sustainability. Both EPA's and GE's fish population/community studies indicate that the PSA contains thriving and self-sustaining populations and communities of fish. Moreover, errors in the interpretation of the Phase I and Phase II studies result in the ERA's conclusion that there is evidence of PCB-related toxicosis, when in fact no consistent relationships between tPCB or TEQ exposure and adverse effects were observed. Consequently, ED50s derived from the Phase II study are flawed. The ERA also misrepresents the EPA and GE field studies, which provide no evidence of harm but are characterized in the risk analysis as showing undetermined evidence of harm. Further, the effects thresholds used to develop site-specific and literature-based MATCs are unduly low due to the use of a number of unnecessarily conservative methods and interpretations of the data. If these issues are addressed, the site-specific toxicity and field studies would support the conclusion that while some effects may be occurring sporadically in individual fish (as indicated in the toxicity studies), they are not affecting the populations; and the literature-based effects thresholds, when recalculated and applied without the unnecessary egg to whole body conversion factor, also would predict low or negligible risks to fish in the PSA.

SUMMARY OF CONCLUSIONS
<u>RISKS TO FISH IN PSA</u>

The data from the site-specific toxicity and field studies, considered together, indicate that, while some effects may be occurring in individual fish, they are not adversely affecting the sustainability or condition of the local fish populations. Thus, there are negligible risks to the overall fish populations and communities in the PSA.

7.5 Extrapolation of Risks to Downstream Reaches

The ERA concludes that there is no risk to warmwater fish downstream of the PSA and that coldwater fish (i.e., trout) are at low risk in Reaches 7 and 9. The ERA characterizes the risks to trout as highly uncertain due to the lack of knowledge regarding the sensitivity of trout species found downstream of the PSA (Vol. 1, p. 5-74; Vol. 5, p. F-77).

As discussed in Section 7.3 and Attachment I, the MATCs used in these comparisons are overly conservative for a variety of reasons. In particular, the MATC developed for trout species, based on dividing the warmwater fish MATC by four, is not justified, both because the warmwater fish MATC is itself is too low and because there is no basis for the application of an arbitrary factor of four. As a result, the ERA overestimates risks to trout in the downstream reaches.

See Table 7-1 for specific comments on Charge Questions 3.3(a)-(j)

Table 7-1. Assessment Endpoint: Survival, Growth and Reproduction of Fish

EPA Charge Question #3.3	GE Response
(a) Were the EPA studies and analyses performed (e.g., field study, site-specific toxicity study, comparison of exposure and effects) appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), and based on accepted scientific practices?	**Phase I and Phase II Toxicity Studies:** Design of studies was appropriate. **EPA Fish Community Study:** Design of study was appropriate. **Fish Tissue Concentrations vs. Literature Effects:** Basic methodology was based on accepted practices, but site-specific effects data should be used when available.
(b) Were the GE studies and analyses performed outside of the framework of the ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), based on accepted scientific practices, and incorporated appropriately in the ERA?	**GE Study on Largemouth Bass Reproduction/Population and Fish Community:** This study was appropriate for evaluating whether the largemouth bass population in the PSA is reproducing and self-sustaining and has population parameters within the expected range. The fish community portion of this study was appropriate for providing an overview of the fish community in the PSA. The ERA's criticisms of this study are unwarranted. The ERA gives the study too low weight and erroneously concludes that it showed "undetermined" evidence of harm when in fact it showed no evidence of harm. (Section 7.2)
(c) Were the estimates of exposure appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?	**Phase I Toxicity Study:** Exposure to individual fish was poorly characterized, since tissue composites (instead of individual fish) were used for the determination of chemical concentrations in adult carcasses. **Phase II Toxicity study:** Exposure is appropriately quantified. **EPA Fish Community Study:** Exposure is not estimated. **Fish Tissue Concentrations vs. Literature Effects:** Exposure is appropriately quantified.
(d) Were the effects metrics that were identified and used appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable)?	**Site-Specific Effects Metrics:** The tissue-based effects metric (MATC) of 49 mg/kg tPCBs for warmwater fish, which is the mean of thresholds from the Phase I (<45 mg/kg) and Phase II (52 mg/kg) studies, is overly conservative. Both the Phase I and Phase II thresholds are based on effects that were not observed consistently in Housatonic River sites (either spatially or temporally), did not show a consistent exposure-response relationship with PCBs, and were not consistent between the Phase I and II studies. In addition, the Phase II effects metric (a) was inappropriately based on pooled endpoints that did not show consistent PCB dose-response relationships, (b) included results for coldwater species, and (c) used an unnecessary egg to whole body conversion factor. (Sections 7.1 and 7.3.1) The tissue-based effects metric of 12 mg/kg tPCBs for coldwater fish, based on dividing the warmwater MATC by 4, is unsupported because (1) the warmwater MATC is too low and (2) there is no basis for the selection of a conversion factor of 4. (Section 7.3.1) **Literature-Based Effects Metrics:** The literature-based effects thresholds are also overly conservative because they: (1) included a highly sensitive species (lake trout) that is not found in the Housatonic; (2) are based (for tPCBs) on an inappropriate averaging of different effects metrics in different tissue types; and (3) used an unnecessary egg to whole body conversion factor. (Section 7.3.2)
(e) Were the statistical techniques used clearly described, appropriate (i.e., objective, consistent, and reasonable), and properly applied for the objectives of the analysis?	**Phase I Toxicity Study:** The statistical analyses conducted were restricted primarily to ANOVAs. This method does not address the lack of exposure-response relationships for most endpoints. (Section 7.1) **Phase II Toxicity Study:** The ED20s and ED50s were inappropriately based on pooled pathology data which do not reflect the lack of consistent dose-response relationships on the individual endpoints. (Section 7.1.1)

Table 7-1. Assessment Endpoint: Survival, Growth and Reproduction of Fish

EPA Charge Question #3.3	GE Response
(f) Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable)?	The ERA states that the risks to fish are low, but improperly characterizes the risks as "ecologically significant." The EPA toxicity studies do not provide consistent evidence of PCB-related toxicosis, and the EPA and GE field studies show thriving and self-sustaining populations and communities of fish in the PSA. Further, the effects metrics developed are overly conservative. Thus, while some effects may be occurring in individual fish (as indicated by the toxicity studies), those individual effects do not appear to be adversely affecting the local populations. (Section 7.4)
(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements could be made.	**Phase I Toxicity Study:** The inconsistent occurrence of effects among sites, and therefore the lack of consistent dose-response relationships, are not discussed. (Section 7.1.1) **Phase II Toxicity Study:** Uncertainties associated with deriving an effects threshold by pooling abnormality data that individually show no consistent dose-response relationships are not recognized. (Sections 7.1.1 and 7.3.1) **Phase I and Phase II Studies:** The lack of concordance between the effects observed between these two phases is not adequately addressed. (Section 7.1) **EPA Fish Community Study:** Uncertainty is adequately addressed. **GE Largemouth Bass Population/Fish Community Study:** Uncertainty is overestimated due to unwarranted criticisms of the study. (Section 7.2) **Fish Tissue Concentrations vs. Literature Effects:** Uncertainty is adequately addressed.
(h) Was the weight of evidence analysis appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable)? If not, how could it be improved?	**Phase I and Phase II Toxicity Studies:** Weight given to these studies (Moderate/High) is appropriate, but the data are not properly interpreted. (Section 7.1) **EPA and GE Fish Community and Population Studies:** Weight given to these studies (Low/Moderate) is too low given that they provide site-specific evidence of self-sustaining populations and communities of fish in the PSA. Also, these studies are erroneously stated to provide "undetermined" evidence of harm, when they in fact showed no evidence of harm. (Section 7.2) **Fish Tissue Concentrations vs. Literature Effects:** Weight given to these comparisons (Moderate) is too high given the site-specific data and uncertainties in the effects thresholds. (Section 7.3.2)
(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?	The ERA overstates risks to trout in downstream reaches due to use of an overly conservative and unsupported MATC for coldwater fish, as noted in response to Question 3.3(d) above. (Section 7.5)
(j) In the Panel members' opinion, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?	**The conclusion that risks to fish in the PSA are "ecologically significant" is not supported by the data.** The data from the site-specific toxicity and field studies, considered together, indicate that, while some effects may be occurring in individual fish, they are not adversely affecting the sustainability or condition of the local fish populations. Thus, there are negligible risks from PCBs to the overall fish populations and communities in the PSA. (Section 7.4)

8. INSECTIVOROUS BIRDS (QUESTION 3.4)

Key Points

- The ERA assesses risks to insectivorous birds based on field studies and model-based HQ analyses of tree swallows and American robins. It concludes that risks to insectivorous birds from exposure to PCBs and TEQs within the PSA are low, but that this conclusion is uncertain due to conflicting outcomes for the field studies (showing no significant effects) and HQs (predicting intermediate to high risks).

- EPA's three-year field study of tree swallows showed little or no evidence of reproductive harm to tree swallows nesting in the PSA.

  - Ø In fact, EPA's tree swallow study supports a site-specific dose-based effects metric of 15.7 mg/kg bw/d, which should be used in the HQ analyses for *other* avian species for which species-specific effects data are not available, as the effects metric for the most tolerant species.

- GE's field study of American robins showed no evidence that exposure to PCBs in the PSA had any adverse effect on the productivity of robins, relative to that of birds in reference areas or as reported in the literature. This study has high statistical power and should be given high weight, rather than the moderate/high weight that is currently assigned in the ERA.

- The HQ analyses presented in the ERA for these species overestimate risks posed by tPCBs by up to 65-fold, due to their reliance on: (a) unnecessary modeling to estimate exposures, when measured data are available; and (b) effects metrics from other studies (and, for robins, based on other species) that should be replaced by species-specific effects metrics derived from the site-specific studies. The undue conservatism of these HQs is demonstrated by the fact that the risks they predict are contradicted by the field data.

- GE's recommended corrections to the HQs would yield more realistic HQ results that generally indicate negligible risks. As such, the HQs would correspond more closely with the results of the field studies, **thereby confirming, with greater certainty, that PCBs and TEQs pose negligible risks to insectivorous birds in the PSA**.

8. INSECTIVOROUS BIRDS (Question 3.4)

The ERA evaluates risks to insectivorous birds using four measurement endpoints. These consist of two site-specific field studies – one on productivity of tree swallows (conducted by EPA contractors) and one on productivity of American robins (conducted by GE contractors) – plus HQs based on modeled exposures and effects for tree swallows and American robins. The ERA concludes that there are low risks to insectivorous birds from exposure to PCBs and TEQs within the PSA, but characterizes that conclusion as uncertain due to conflicting outcomes for the field studies and the HQs (Vol. 2, p. 7-92). GE has identified a number of specific problems in the ERA's interpretation of the underlying studies, assumptions used in the HQs, and application of the weight-of-evidence approach. Correction of these problems would yield greater concordance among the measurement endpoints and confirm, with greater certainty, that there is negligible risk to insectivorous birds. Our chief concerns are discussed below, while specific comments related to Questions 3.4(a)-(j) are provided in Table 8-1 (at the end of this section).

8.1 EPA's Site-Specific Field Study of Tree Swallows

From 1998 through 2000, EPA contractors (Custer 2002) conducted a field study on productivity of tree swallows. The study monitored nest box occupancy rate, clutch size, daily egg survival, and concentrations of tPCBs and TEQs in pippers (eggs and hatchlings) and 12-day-old nestlings. The ERA concludes that this study showed little or no evidence of harm to tree swallows nesting in the PSA (Vol. 2, p. 7-82, Tables 7.5-4, 7.5-5).[16] GE generally agrees with this interpretation of the tree swallow nest box study.

In fact, the data from this study should be used to develop a site-specific dose-based effects metric to use in the HQ analyses of *other* avian species in the ERA for which species-specific effects data are not available. For those other species, the ERA currently uses a literature-derived effects metric of 7.0 mg/kg body weight/day (bw/d), based on minor effects in American kestrels (from Fernie et al. 2001), as the effects metric representative of the most tolerant avian species (see Vol. 2, pp. 7-69, 7-74, 8-34, 8-35). However, the average of the mean modeled exposure levels reported in the ERA for the three locations in

---

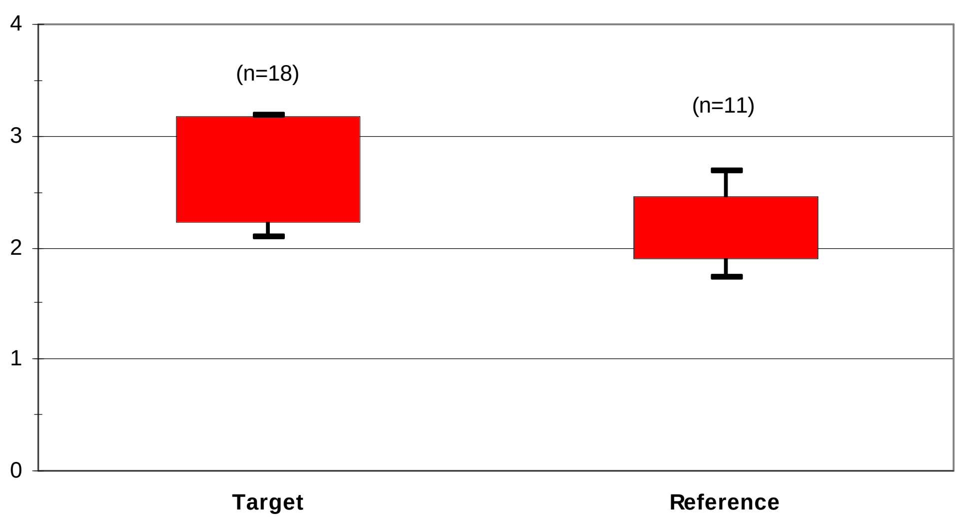
[16] The ERA notes that, in this study, tPCBs were negatively but weakly correlated with hatching success (based on clutches with dead embryos) in 1998 and 1999, but not in 2000, and it speculates that the latter finding may be due to reduced concentrations and cold weather in 2000 (Vol. 2, p. 7-78). However, only one period of cold weather occurred during the incubation period. Additionally, elevated mortality in 2000 only occurred at one (reference) site, rather than at all sites, as would be expected if weather were the cause. Furthermore, 2000 temperature data reported in the local newspaper (the *Berkshire Eagle*) reflected patterns similar to those recorded for 1998. For these reasons, the results from 2000 should not be discounted on account of weather.

the tree swallow study where minor effects were found was 15.7 mg/kg bw/d (Vol. 5, p. G-20, Table G.2-9). As discussed further below (Sections 8.3.2, 9.3.2, 12.3), this value should replace the kestrel-based effects metric of 7.0 mg/kg bw/d as the metric representative of the most tolerant avian species, since Custer (2002) demonstrated that tree swallows are more than twice as tolerant of PCBs as are American kestrels.

8.2 GE's Site-Specific Field Study of Robins

Dr. Scott Robinson (University of Illinois at Urbana-Champaign) and ARCADIS were retained by GE to study American robins breeding in 106 natural nests in the Housatonic River watershed in 2001 to determine whether predation, Mayfield nest success, clutch size, viability, incubation period, hatching, nestling period, or fledging differed significantly between the exposed population and the reference population (Henning et al. 2002).[17] A manuscript on this study, which will be published in the October 2003 issue of *Environmental Toxicology and Chemistry*, is provided as Attachment K of these Comments. This study showed no adverse effects of PCB exposure on any of the endpoints evaluated – a conclusion with which the ERA concurs (even with EPA's reanalysis of the data) (Vol. 2, pp. 7-72, 7-79). Figure 8-1 below illustrates the absence of any differences in fledging success between target (exposed) and reference robins, despite substantial differences in exposure (Figure 8-2).

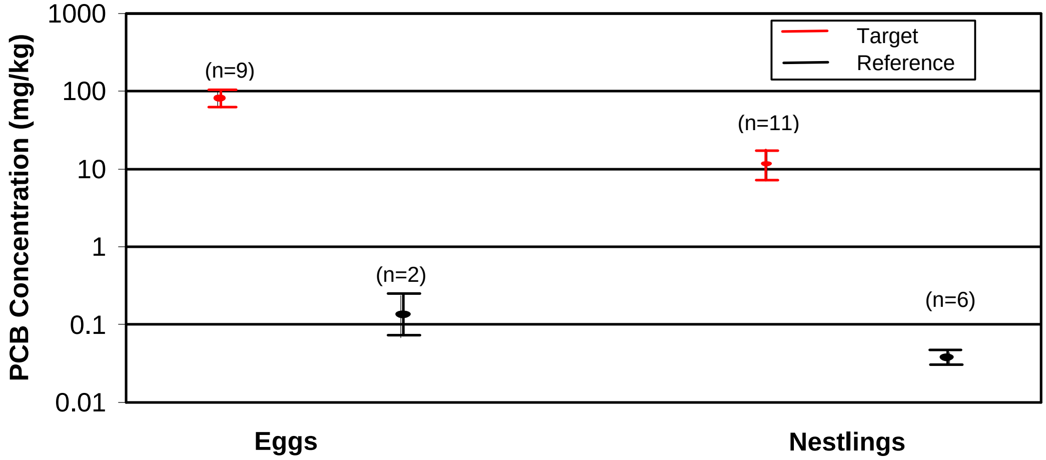
Figure 8-1. Number of Robins Fledged per Nest



Note: Values shown are range-low number fledged (-1 standard error) and range-high number fledged (+1 standard error), as described in Attachment K

[17] This represents the complete list of endpoints evaluated, in contrast with the abbreviated list described in the ERA (Vol. 2, p. 7-71; Vol. 5, p. G-68). The full report on this study (Henning et al. 2002) is provided as part of the compact disk containing most of the GE-sponsored studies.

Figure 8-2. PCB Exposure in Target and Reference Robins



Note: Values shown are mean +/- 1 standard error

The ERA raises a number of concerns about this study which cause it to give the study an overall weight of only Moderate/High, rather than the High weight given to the tree swallow field study (Vol. 2, p. 7-79, Table 7.5-3; Vol. 5, pp. G-69 - G-77). Those concerns are unfounded. One key error in the ERA's assessment is discussed below, while the others are addressed in Attachment A.

The ERA attributes uncertainty in the robin field study to the small number of samples analyzed for PCBs, which supposedly limited the study's power to detect differences in reproductive success (Vol. 2, p. 7-79; Vol. 5, pp. G-74, G-77). While sample sizes for the egg and nestling tissue analyses were indeed small, these data were used primarily to demonstrate that exposed and reference populations were distinct. These data show significant differences in concentrations ($p < 0.05$) with high statistical power (0.76 to 0.96) (Henning et al. 2002, Table 3). Because the two populations were distinct, it was appropriate to use statistical analyses to compare reproductive measures in exposed nests ($n \leq 62$) and reference nests ($n \leq 44$). Both power analyses and bioequivalence testing are presented in the study report (Henning et al. 2002), illustrating that the sample sizes are sufficient to show differences (or lack thereof) in productivity between the exposed and reference populations.

8.3 Modeled Exposures and Effects

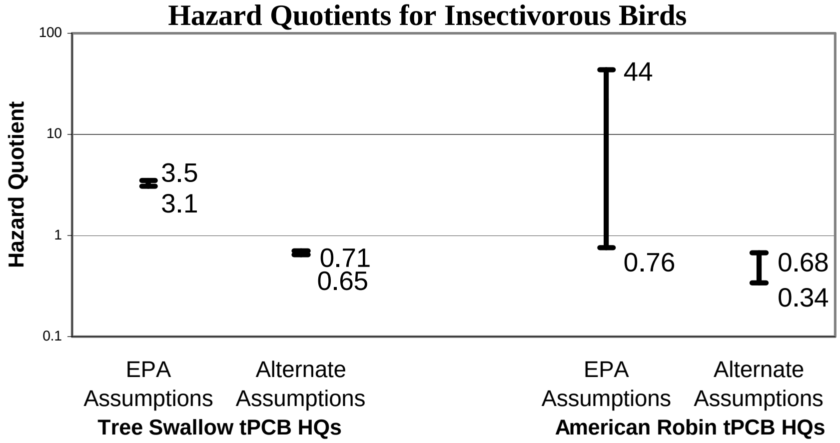
The ERA also includes HQs based on modeled exposures and effects to evaluate potential risks to tree swallows and American robins, concluding that these lines of evidence show both species to be at high risks in the PSA due to tPCBs (Vol. 2, p. 7-82). However, there are several problems with the HQs, which cause them to substantially overestimate risk.

8.3.1 Modeled exposures and effects for tree swallows

The tree swallow HQs for tPCBs compare modeled concentrations of tPCBs in nestlings to a tissue-based tPCB effects threshold range of 62.2 mg/kg (based on nestling effects from McCarty and Secord 1999) to 69 mg/kg (based on pipper effects from Custer 2002) (Vol. 2, pp. 7-74, 7-75; Vol. 5 pp. G-58, G-63 - G-64). First, the use of McCarty and Secord (1999) contributes unnecessary uncertainty to the lower-bound effect metric. Instead, Custer's (2002) threshold of 62.8 mg/kg for hatching problems in 1998 (Vol. 2, p. 7-71; Vol. 5, p. G-46) should be used as a lower-bound effect metric, given its site-specificity and superior study design. Although that value is similar to the 62.2 mg/kg effect metric generated from McCarty and Secord (1999), the McCarty and Secord (1999) study is not site-specific, does not report the age of nestlings collected, and has numerous other problems (detailed in the ERA, Vol. 2, p. 7-88; Vol. 5, p. G-57). More importantly, the HQ comparisons should not be based on modeled tissue concentrations for 15-day-old nestlings, since the effects metrics from Custer (2002) were developed for hatching success based on pipper (egg and hatchling) tissue concentrations. Instead, actual measured data on tPCB concentrations in the correct age class – pippers (for which over 200 measurements are available [Vol. 5, Table G.2-22]) – should be used to represent exposure in this HQ.

These recommended changes would result in an approximate five-fold reduction in the value of the tPCB HQs. For example, the median modeled tPCB tissue concentration for 15-day-old nestlings at Holmes Road (the site of maximum exposure) was 215 mg/kg (Table G.2-13); dividing that value by the ERA's lower-bound effects metric of 62.2 mg/kg yields an HQ of 3.5. However, when the median measured pipper concentration (44.9 mg/kg; Table G.2-22) for the same location is used to characterize exposure and the recommended lower-bound effects metric of 62.8 mg/kg tPCBs is used, the HQ is reduced to 0.71. The impact of the recommended changes to the exposure and effects assumptions for the tree swallow tPCB HQs is illustrated in Figure 8-3.

Figure 8-3. Impact of Alternative Assumptions on tPCB Hazard Quotients for Insectivorous Birds



Note: Values shown are the upper and lower bound HQs

The HQs for TEQs are also overly conservative, because an inappropriate study (Nosek et al. 1992) provides the basis for the lower-bound effects metric (44 ng/kg bw/d). Nosek et al. (1992) employed a dosing regimen consisting of weekly TCDD intraperitoneal injections in a corn oil and acetone vehicle. Such a dosing regime cannot accurately represent dietary exposures to TCDD in the wild due to accelerated absorption, enhanced storage, and the resultant acute toxic injury. Daily doses from food or water better simulate exposure in the wild and are more easily metabolized and excreted, thus reducing the overall body burden and toxicity. Therefore, it is not appropriate to use the results from the Nosek et al. (1992) study to derive an acceptable daily dose from TCDD in sensitive bird species.

The study by Hoffman et al. (1996), which serves as the basis for the upper-bound TEQ effects metric in the ERA (Vol. 2, p. 7-74; Vol. 5, p. G-61), is the only study identified that provides an appropriate basis for estimating a TEQ toxicity threshold in sensitive avian species. All other avian toxicity studies for TEQs that are cited in the ERA are egg injection studies, which also cannot adequately simulate dietary exposure in the wild. As discussed in the ERA, Hoffman et al. (1996) supports a toxicity threshold 25,000 ng/kg bw/d (Vol. 2, p. 7-70; Vol. 5, p. G-60), which should serve as the sole avian effects metric for TEQs. Use of this effects metric would result in more than a 500-fold reduction in the TEQ-based HQs for tree swallows.

8.3.2 Modeled exposures and effects for American robins

The robin HQs for tPCBs were calculated by comparing modeled doses of tPCBs in breeding robins to a threshold dose range reportedly representing the most sensitive avian species (white leghorn chickens) and the most tolerant species (American kestrels) (Vol. 2, p. 7-74; Vol. 5, p. G-59). The most important limitation of the robin HQs is the use of a range of tPCB effects metrics based on white leghorn chickens (Lillie et al. 1974) and American kestrels (Fernie et al. 2001). The robin study itself (Attachment K) presents a NOAEL of 7.8 mg/kg bw/d. Because that study was species-specific, site-specific, and stressor-specific, this NOAEL is an appropriate effects metric for robins. However, some uncertainty is associated with this dose, because it was estimated, rather than measured. Therefore, GE recommends pairing this value with the dose-based effects metric generated from the Custer (2002) field study on tree swallows – i.e., 15.7 mg/kg bw/day, as discussed above. Tree swallows are suitable surrogates for robins, in that they are relatively similar in size and share the same taxonomic order and feeding guild. Furthermore, the dose-based effects metric from Custer (2002) is site-specific and stressor-specific.

It should also be noted that the robin HQs use a modeled food intake rate developed from an algorithm that is based on passerines (perching birds) in general, rather than robins in particular, despite the availability of more accurate measured rates for both free-living robins (Skorupa and Hothem 1985) and captive robins (Hazelton et al. 1984).[18] These measured rates should be used.

The recommended change to the tPCB effects metrics for robins would result in up to a 65-fold reduction in the HQs. For example, the ERA estimates that the median tPCB dose at Site 13 (the site of maximum exposure) was 5.3 mg/kg bw/d (Table G.2-29); dividing that value by the ERA's lower-bound effects metric of 0.12 mg/kg bw/d yields an HQ of 44. However, if the same measure of exposure (5.3 mg/kg bw/d) is divided by the recommended lower-bound effects metric of 7.8 mg/kg bw/d based on the robin field study, the HQ is reduced to 0.68. The impact of the recommended changes to the exposure and effects assumptions for the American robin tPCB HQs is illustrated in Figure 8-3 above.

Finally, as discussed above for tree swallows, the selection of effects metrics for TEQs (Vol. 2, p. 7-74; Vol. 5, pp. G-60 - G-61) inappropriately includes a study based on weekly intraperitoneal injections (Nosek et al. 1992). The only available study that provides a defensible TEQ effects metrics for birds is the Hoffman et al. (1996) study, which yielded an effects metric of 25,000 ng/kg bw/d. Use of this effects metric would result in more than a 500-fold reduction in the TEQ-based HQs for robins.

8.4 Overall Assessment

The ERA concludes that the HQs support a finding of intermediate to high risks for tree swallows and American robins in the PSA, but that the more highly weighted field studies suggest that, if effects are occurring, they are minor (Vol. 2, p. 7-90; Vol. 5, p. G-90). Thus, the ERA's weight-of-evidence evaluation concludes that risks are low, but it qualifies this conclusion as uncertain due to conflicting results for the different lines of evidence. For the reasons given above, GE believes that: (1) the robin field study warrants an overall weighting of High, rather than Moderate/High; and (2) corrections should be made to the HQ analyses for these species, both in the exposure estimates and in the effects metrics used. With these recommended corrections, HQs would indicate negligible risk. Consequently, the various lines of evidence would more closely agree and would, with greater certainty, confirm that insectivorous birds are at negligible risk from exposure to PCBs and TEQs in the PSA.

---

[18] The ERA incorrectly asserts (Vol. 2, p. 7-22) that the only published study on food intake rate for robins is based on captive robins. As cited in the *Wildlife Exposure Factors Handbook* (EPA 1993), a study on wild robins (Skorupa and Hothem 1985) is also available.

SUMMARY OF CONCLUSIONS
<u>RISKS TO INSECTIVOROUS BIRDS</u>

The strong field studies showing no PCB effects on tree swallows and American robins, together with the correction of the HQ analyses, indicate that insectivorous birds are at negligible risk from exposure to tPCBs and TEQs in the PSA.

See Table 8-1 for specific comments on Charge Questions 3.4(a)-(j)

Table 8-1. Assessment Endpoint: Survival, Growth, and Reproduction of Insectivorous Birds

EPA Charge Question #3.9	GE Response
(a) Were the EPA studies and analyses performed (e.g., field study, site-specific toxicity study, comparison of exposure and effects) appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), and based on accepted scientific practices?	**Tree Swallow Field Study:** Methodology appears sound. **Tree Swallow and Robin HQs:** Basic methodology is appropriate, but overly conservative assumptions were used for both exposure and effects (see below).
(b) Were the GE studies and analyses performed outside of the framework of the ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), based on accepted scientific practices, and incorporated appropriately in the ERA?	**Robin Field Study:** This study used accepted methodology and results have high power. The ERA's interpretation of the overall conclusions of this study is appropriate, but the ERA advances some unwarranted criticisms and assigns the study too low weight (should be High, rather than Moderate/High). (Section 8.2)
(c) Were the estimates of exposure appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?	**Tree Swallow Field Study:** Exposure is appropriately quantified. **Robin Field Study:** Exposure is appropriately quantified; the ERA's criticisms of the methods used to characterize exposure are unfounded. (Section 8.2 and Attachment A). **Tree Swallow HQs:** The tissue-based exposure metric is modeled for an age class (15-day nestlings) that differs from that used in the effects metric. Site-specific pipper measurements would be a more appropriate measure of exposure. (Section 8.3.1) **Robin HQs:** The estimated food intake rate is modeled from an algorithm for birds in general, rather than for robins in particular. A more accurate measured rate is available for free-living robins and should be used. (Section 8.3.2)
(d) Were the effects metrics that were identified and used appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable)?	**Tree Swallow Field Study:** Effects assessment was generally conducted appropriately. However, data from 2000 should not be discounted due to weather. (Section 8.1 and note 16) **Robin Field Study:** Effects results presented in original report and manuscript (Attachment K) are correct and should be presented in the ERA's Effects Assessment. **Tree Swallow HQs:** For tPCBs, the lower-bound tissue-based effects metric is based on a non-site-specific study that has numerous design and execution problems. Instead, the lower-bound effects metric for tPCBs should be based on the threshold for hatching problems in 1998 from the site-specific tree swallow study. (Section 8.3.1) For TEQs, the lower-bound dose-based effects metric is inappropriately based on a study that used weekly intraperitoneal injections, which do not accurately simulate dietary exposures in the wild. Only one study (Hoffman et al. 1996) offers a defensible basis for an avian TEQ effects metric and should serve as the basis for the TEQ effects metric for tree swallows. (Section 8.3.1). **Robin HQs:** For tPCBs, effects metrics for robins should be based on those derived from the site-specific studies on robins and tree swallows, instead of the ERA's overly conservative values for the most sensitive and tolerant avian species. For TEQs, the effects metric derived from Hoffman et al. (1996) should serve as the only TEQ effects metric, because no other studies provide a defensible basis for one. (Section 8.3.2)

Table 8-1. Assessment Endpoint: Survival, Growth, and Reproduction of Insectivorous Birds

EPA Charge Question #3.9	GE Response
(e) Were the statistical techniques used clearly described, appropriate (i.e., objective, consistent, and reasonable), and properly applied for the objectives of the analysis?	**Tree Swallow Field Study:** Statistical techniques applied in the ERA appear correct. **Robin Field Study:** Statistical analyses presented in original report and manuscript (Attachment K) on this study are appropriate and should be presented in the ERA. ERA's reanalysis of the data from this study is flawed and lacks transparency. (Section 8.2 and Attachment A) **HQs:** Mathematical calculations appear correct.
(f) Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable)?	**Tree Swallow Field Study:** Overall risk characterization is appropriate. **Robin Field Study:** Overall risk characterization is appropriate. **Tree Swallow HQs:** Risks from tPCBs are overestimated by > 5 times due to use of inappropriate exposure concentrations. Risks from TEQs are overestimated by > 500 times due to use of an unsupported effects metric. (Section 8.3.1) **Robin HQs:** Risks from tPCBs are overestimated by up to 65 times due to use of overly conservative exposure and effects assumptions. Risks from TEQs are overestimated by > 500 times due to use of an unsupported effects metric. (Section 8.3.2)
(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements could be made.	**Tree Swallow Field Study:** Uncertainty analysis appears generally appropriate. **Robin Field Study:** Uncertainties are overstated. (Section 8.2 and Attachment A) **HQs:** Impacts of overly conservative assumptions are not adequately recognized. Improvements are recommended in Section 8.3.
(h) Was the weight of evidence analysis appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable)? If not, how could it be improved??	**Tree Swallow Field Study:** Weight given to this study (High) is appropriate. **Robin Field Study:** Because criticisms of this study are unfounded, the weight assigned (Moderate/High) is too low. The study warrants a High weight. (Sections 8.2 and 8.4) **HQs:** Due to uncertainties and over-conservatism in assumptions, weight given to HQs (Moderate) is too high. (Section 8.3)
(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?	Not applicable.
(j) In the Panel members' opinion, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?	**The ERA's conclusion that the risks to insectivorous birds in the PSA are low but uncertain overstates the uncertainty.** The strong field studies showed no PCB effects on tree swallows or robins. Corrections to the HQs would likewise show negligible risks to these species, thereby confirming with greater certainty that PCBs and TEQs pose negligible risks to insectivorous birds in the PSA. (Section 8.4)

9. PISCIVOROUS BIRDS (QUESTION 3.5)

Key Points

- The ERA assesses risk to piscivorous birds based on a field study of belted kingfishers and modeled exposures and effects (HQs) for belted kingfishers and ospreys.

  - Ø It characterizes risks to belted kingfishers as low, with uncertainty due to inconsistent results between the field study and the HQ analysis.

  - Ø It characterizes risks to ospreys as high for tPCBs (and intermediate for TEQs) based on the HQ analyses, with uncertainty due to the availability of only one line of evidence.

- The field study of belted kingfishers – the most highly exposed piscivorous bird known to breed in the PSA – found no evidence of harm. The local population is breeding successfully and the density of the population is consistent with the quality of the available habitat.

- Ospreys are not a relevant representative receptor, because they do not breed in western Massachusetts. Because kingfishers have higher exposure potential than any other piscivorous bird species known to breed within the PSA, they suffice as the single representative of this feeding guild.

- If, however, a second representative piscivorous bird species is required, then great blue herons would be a more appropriate choice than ospreys. Great blue herons do breed within foraging range of the river, and field data are available on the productivity of great blue herons breeding within foraging distance of the river versus those breeding elsewhere in Massachusetts. These data could serve as a second line of evidence for this species, in addition to an appropriate HQ analysis.

- The HQs substantially overestimate risks to both belted kingfishers and ospreys as a result of: (a) overly conservative assumptions related to food intake rate and (for ospreys) foraging time within the PSA; and (b) use of inappropriate effects metrics for chickens and kestrels, when more supportable and relevant effects metrics are available. The prediction of risks to belted kingfishers based on the HQs, when the site-specific field study shows no evidence of harm, demonstrates the excessive conservatism of the HQs.

- **A more balanced assessment of the evidence, including use of belted kingfishers alone or belted kingfishers and great blue herons as representative species, indicates that piscivorous birds are likely at negligible or low risk in the PSA.**

  - Ø Although ospreys should not be used as a representative species, correction of the osprey HQs to apply to migrating ospreys would result in HQs below 1, thus supporting a similar conclusion.

9. PISCIVOROUS BIRDS (Question 3.5)

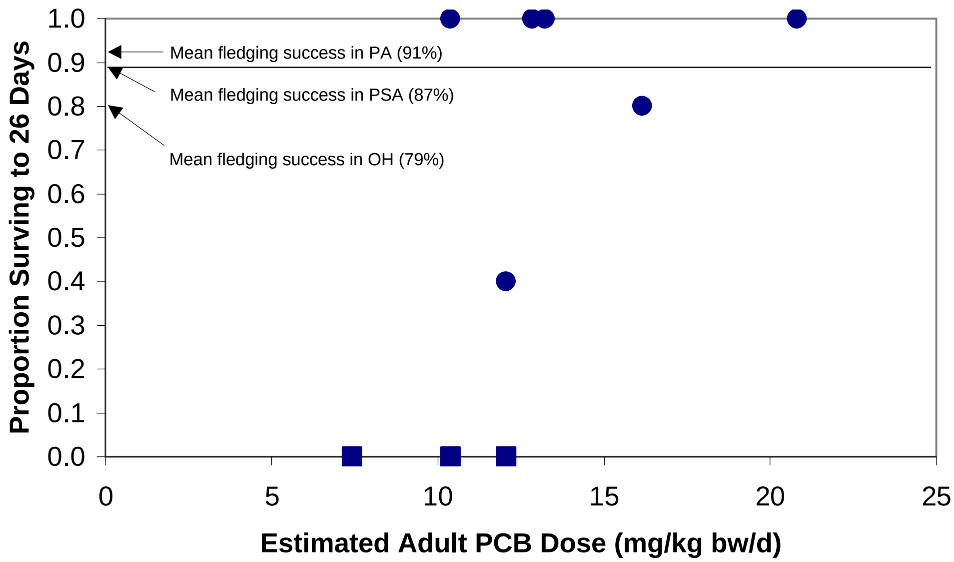
The ERA evaluates risks to piscivorous birds based on two types of measurement endpoints: (1) a field study of productivity of belted kingfishers breeding within the PSA, conducted by GE contractors; and (2) modeled exposures and effects (i.e., HQs) to evaluate risks posed to belted kingfishers and ospreys. The ERA concludes that belted kingfishers in the PSA are at low risk, although that conclusion is qualified as uncertain because the field study and the HQ analysis did not provide concordant results (Vol. 2, p. 8-48 - 8-49). The ERA concludes that ospreys in the PSA are at high risk, although that conclusion is qualified as uncertain because only one line of evidence is available (Vol. 2, p. 8-49). GE's principal concerns with these conclusions relate to the inappropriate selection of ospreys as a representative species for the PSA and the use of several unnecessary and overly conservative assumptions in the HQ analyses for both kingfishers and ospreys. Our chief concerns are discussed below, while specific comments related to Questions 3.5(a)-(j) are provided in Table 9-1 (at the end of this section).

9.1 GE's Site-Specific Field Study on Kingfishers

In 2002, ARCADIS and Dr. Robert Brooks of Pennsylvania State University, at GE's request, conducted a study of belted kingfishers breeding along the Housatonic River to determine whether productivity or population density is impaired, in light of the presence of PCBs in their prey. The full report on this study is provided as Attachment L of these Comments. The scope of the study involved: (1) characterizing habitat quality throughout the PSA using the U.S. Fish and Wildlife Service's Habitat Suitability Index (HSI) model (Prose 1985); (2) monitoring kingfisher burrows from courtship through fledging; (3) collecting nest remnants to characterize predominant prey; (4) estimating tPCB doses based on available biota chemistry data; (5) testing whether the estimated dose of tPCBs, habitat suitability, fledge date, or nest density were significant predictors of survival to 26 days; and (6) comparing observations to data reported in the scientific literature.

Ten kingfisher burrows were identified along the Housatonic River, one of which was subsequently destroyed by human excavation and three of which were later depredated. The study concluded that the study area supports a kingfisher population consistent with the quality of the available habitat, and that that population is breeding successfully, with productivity consistent with values reported by Dr. Brooks for sites in Ohio and Pennsylvania with similar habitat (Brooks and Davis 1987). It also showed no evidence of a relationship between productivity and estimated tPCB dose, as illustrated on Figure 9-1.

Figure 9-1. Survival of Belted Kingfisher Nestlings



Note: Squares = depredated nests in PSA; circles = successful nests in PSA; mean values exclude depredated nests

While the ERA agrees with the overall conclusion of the belted kingfisher field study – namely, that it provides no evidence of adverse effects on reproduction and confirms the presence of successfully breeding kingfishers in the PSA – it raises a number of concerns with the study that we believe are overstated. A few examples are addressed below, while others are addressed in Attachment A.

- The ERA notes that the sample size was small (Vol. 2, p. 8-41; Vol. 6, p. H-49). However, given the size of the PSA and the relatively poor foraging habitat available within the PSA, the nine pairs monitored likely represent reasonably full usage of the PSA. Furthermore, the ability of the study to detect an expected effect of phenology on fledging success (i.e., the proportion of nestlings surviving to 26 days was significantly higher with earlier fledge dates) suggests that it was adequately sensitive to detect differences among nests in fledging success, even with the low sample size.

- The ERA states that uncertainty results from the method used to estimate dose to the kingfishers, which was based on tPCB concentrations in prey within foraging distance from their burrows (Vol. 2, pp. 8-40, 8-47; Vol. 6, pp. H-52, H-61). While the prey sampling locations were not precisely known, they were classified based on the reach of river extending from one-half mile upstream to one-half mile downstream of the river mile specified for a given sample. Because all prey sampling locations were either well within foraging distance of a burrow or well beyond twice the foraging distance of a burrow, assignment of samples to burrows was unambiguous. Hence, the method used to estimate dose to the kingfishers contributes less uncertainty than is suggested in the ERA.

- The ERA asserts that the intake calculations presented as part of the kingfisher study yield a dose gradient that is too narrow to evaluate a dose-response relationship (Vol. 5, p. H-47). The gradient reported in the kingfisher study – 7.4 to 21 mg/kg bw/d – varies by approximately three-fold. The ERA's calculation of minimum and maximum tPCB doses within the PSA also ranges by approximately three-fold, from 10.1 to 30.4 mg/kg bw/d (Vol. 5, Table H.2-6). Hence, the dose range reported by Henning et al. (2002) is not an artifact of a poor study design, as is implied in the ERA; rather it reflects the actual site conditions.

9.2 Selection of Ospreys as Representative Receptor

The ERA evaluates potential risks to breeding ospreys based only on an HQ. However, the selection of breeding ospreys as a "representative" receptor is inappropriate, because the possibility of this species breeding in western Massachusetts is entirely hypothetical and is not supported by the field surveys. The ecological characterization (ERA, Vol. 3, pp. 5-9 - 5-10) and the ERA (Vol. 6, pp. H-2, H-17) confirm the absence of any observation of breeding ospreys in the PSA during three years of intensive field work. However, the ERA includes breeding ospreys as representative piscivorous birds based on (1) the assumption that, as the regional population expands, ospreys may nest in the PSA in the future (Vol. 6, p. H-2); (2) the hypothesis that ospreys are not currently breeding within the PSA because of the presence of contaminants (Vol. 2, p. 8-49). The first assumption is wholly without scientific support. The second hypothesis is likewise unwarranted; there is no evidence that osprey breed *any place* in western Massachusetts, including areas with no PCBs. Alternative recommendations regarding representative species are presented below.

During three seasons of intensive field activities on the PSA, ospreys were observed on only six occasions (Vol. 3, Section III, pp. 5-9 - 5-10). All observations occurred in late summer or early fall, concurrent with fall migration period for ospreys. The ecological characterization appropriately concludes that all observations were of transient (i.e., migratory) individuals, rather than breeding individuals. Several independent sources support the conclusion that any ospreys observed in this region are likely transients. A press release from the Massachusetts Division of Fish and Wildlife (MDFW) (http://www.state.ma.us/dfwele/Press/prs9708.htm) notes that the westernmost osprey nest in the state in 1997 was located in Westborough (more than 80 miles east of Pittsfield). The Massachusetts Audubon Society reported that the only osprey pairs breeding inland in Massachusetts in 2002 were located in Pepperell (approximately 85 miles east of Pittsfield) and again in Westborough (pers. comm. with Wayne Petersen, Sept. 11, 2003). The Breeding Bird Survey, a nationwide annual survey of breeding populations throughout the United States conducted since 1966 and overseen by the U.S. Geological Survey, has *never* recorded ospreys breeding in western Massachusetts (www.mbr-pwrc.usgs.gov). Christmas Count surveys, which are overseen by the National Audubon Society, indicate that ospreys do not overwinter in Massachusetts, as there are also no records of osprey (www.audubon.org/bird/cbc). *Birds of Massachusetts* (Veit and Petersen 1993), a comprehensive breeding bird atlas for the state, indicates that ospreys breed in Massachusetts only along the Atlantic Coast. That atlas also notes that fall migrations over western Massachusetts occur in high numbers, with a maximum of 66 individuals recorded at Mount Tom in Holyoke on September 20, 1963. Based on these independent sources, as well as the ecological characterization conducted for the ERA, there is no evidence that supports the ERA's assumption that ospreys are likely to breed at the PSA in the future.

- **Use of ospreys as a representative receptor is inappropriate because ospreys do not breed in western Massachusetts.**

- **The ERA acknowledges: "In general, ospreys were uncommon visitors to the study area, and no active nesting was observed. Most observations were during late summer and fall (i.e., the fall migration period)." (Vol. 3, Species Profile: Osprey, p. 5)**

The ERA also speculates that the presence of contaminants has caused ospreys not to breed at the PSA (Vol. 2, p. 8-49). Based on the evidence, discussed above, indicating that the PSA is outside the breeding range of ospreys, there is no support for that speculation. Moreover, if it were the case that contamination has lead to breeding failure in ospreys, one would expect to find records of failed breeding attempts by ospreys within the PSA. However, because there are no records of ospreys in the PSA during the breeding season, it follows that there are also no records of either successful or failed breeding attempts by ospreys in the PSA. We were also unable to identify any records of breeding attempts by ospreys at other large waterbodies in western Massachusetts. In short, regardless of the suitability of the PSA habitat for ospreys, the PSA lies outside of the natural breeding range for ospreys, and that observation is a far more likely explanation for the absence of breeding ospreys than the ERA's conjecture regarding contamination.

For all these reasons, breeding ospreys should not be included as a representative piscivorous bird species in the ERA. The endpoint should focus only on the most highly exposed piscivorous bird species known to breed within the PSA. Based on their relatively small body size, high foraging time, and aquatic diet, belted kingfishers are expected to be the most highly exposed piscivorous bird species known to breed within the PSA. It follows that belted kingfishers are an appropriate and conservative surrogate species for other piscivorous birds that may visit the PSA.[19] Moreover, two lines of evidence are available for kingfishers.

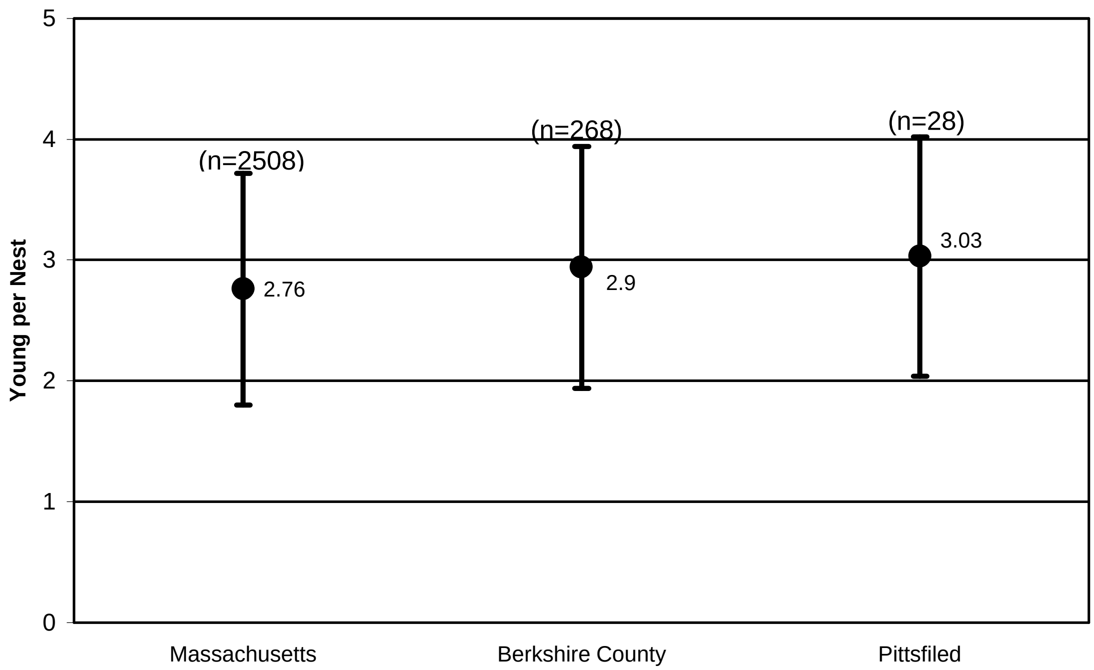
If a second representative of the feeding guild is critical, however, great blue herons would be a more appropriate choice than ospreys. In contrast to ospreys, great blue herons are known to breed within foraging range of the PSA. For nearly two decades, MDFW has collected productivity data (i.e., young

[19] The ERA indicates that belted kingfishers are not a suitable surrogate for ospreys due to physiological differences between the two species (Vol. 6, p. H-58). However, the ERA evaluates risks to hypothetical breeding ospreys based on effects reported for white leghorn chickens and American kestrels, which also differ physiologically from ospreys, as well as with respect to feeding guild and (in the case of chickens) domestication. In any case, it is illogical to reject belted kingfishers on the basis that they are not a suitable surrogate for a species that does not breed in the PSA in any event.

per nest) for great blue herons breeding throughout Massachusetts (MDFW 1979, 1980, 1981, 1982, 1983, 1984, 1985, 1986a,b, 1987, 1989, 1991, 1996), including numerous sites within foraging range of the PSA. Following the same methodology, GE contractors completed surveys in Berkshire County in 1995, 1997, 1998, and 1999, including at a colony in Pittsfield that is closest to the Housatonic River.

The data resulting from both MDFW and GE surveys are presented in Attachment M of these Comments and are summarized in Figure 9-2 below. As shown in Attachment M, statistical analyses found no evidence that numbers of young per nest were lower in either Berkshire County or the Pittsfield colony (either of which may be assumed to define foraging distance from the Housatonic River, depending on the assumed foraging range of great blue herons) than in the other colonies.

Figure 9-2. Great Blue Heron Productivity in Massachusetts (1980-1999)



Note: Values shown are mean +/- 1 standard deviation

Although measurements of exposure are not available for these great blue herons, the productivity data reflect the actual performance of the existing populations most likely to forage in the Housatonic River. Indeed, the colony in Pittsfield is located only two miles from the PSA. Hence, these field productivity data provide a second line of evidence that could be used, in addition to an appropriately conducted HQ analysis, to assess risks to great blue herons. Thus, compared to ospreys, great blue herons offer a more meaningful and defensible representative of piscivorous bird species.

9.3 Modeled Exposures and Effects

The ERA employs HQs to evaluate potential risks to belted kingfishers and ospreys, concluding that these HQs show both species to be at high risk from tPCBs and intermediate risk from TEQs (Vol. 2, Tables 8.5-3, 8.5-4). HQs for both species were calculated by comparing modeled doses to a range of literature-based effects metrics based on the most sensitive avian species and a tolerant avian species (Vol. 2, p. 8-35; Vol. 6, p. H-40). As detailed below, there are several problems with the exposure assumptions and effects metrics used in these HQs, which result in substantial overestimates of risk and unwarranted uncertainty.

9.3.1 Exposure assumptions

The HQs for both belted kingfishers and ospreys rely on the use of a modeled food intake rate developed from a bird algorithm that is not specific to these species. As previously discussed (Section 4.2), use of this general algorithm requires inputs for various factors for which limited data are available, and the results are thus highly uncertain (as shown by the strong influence of these variables in the sensitivity analyses [Vol. 6, Tables H.2-8 & H.2-16]).[20] However, this approach is unnecessary since EPA's *Wildlife Factors Handbook* provides measured food intake rates for free-living kingfishers (Alexander 1977) and ospreys (Poole 1983). The excess conservatism resulting from the ERA's use of the modeled food intake rate rather than the measured values is illustrated by the finding that the field-based estimate for free-living kingfishers reported by Alexander (1977) is close to the 30th percentile of the modeled food intake rate used in the ERA (Vol. 2, p. 8-14; Vol. 6, p. H-12). We believe that the measured food intake rates should be used for both species, as recommended in EPA's *Wildlife Exposure Factors Handbook.*

A second overly conservative exposure assumption for ospreys relates to their assumed foraging time, which is defined as 1.0 (i.e., ospreys are assumed to derive 100% of their prey from the PSA). As discussed in detail above, the only ospreys documented at the PSA were migratory. Hence, the foraging time should reflect the likelihood that ospreys only forage at the PSA for a few days, while passing through during migration. A more realistic foraging time would be 0.3% to 0.8% (i.e., 1/365 to 3/365).

---

[20] Conservatism in the modeled food intake rate for ospreys is exacerbated through the ERA's use of a model developed for Charadriiformes, because birds in this order tend to be substantially smaller than ospreys, and hence their metabolism and food intake rate are substantially higher than those of ospreys. Food intake rates for all other avian receptors were modeled based on the general equation for birds. If the food intake rate for ospreys had been modeled using the general equation for birds instead of the Charadriiformes equation, the food intake rate would have been reduced by approximately 50%.

9.3.2 Effects metrics

In the absence of toxicity studies providing suitable dose-response data for belted kingfishers or ospreys, the ERA employs a threshold range for reproductive effects in the most sensitive and a tolerant avian species. For tPCBs, that range is defined as 0.12 mg/kg bw/d, based on Lillie et al.'s (1974) study on white leghorn chickens, to 7.0 mg/kg bw/d, based on Fernie et al.'s (2001) study on American kestrels (Vol. 2, p. 8-35; Vol. 6, p. H-40). GE believes that both of these values are unsupported.

For the upper bound of the range, the ERA's reliance on the study of American kestrels to represent the most tolerant avian species is inappropriate, because, as discussed above and acknowledged in the ERA itself, "the most tolerant bird species to tPCBs found in the literature was the tree swallow" (Vol. 2, p. 7-90). Moreover, the Custer (2002) study provides the basis for a site-specific and stressor-specific dose-based effect metric for the tree swallow – 15.7 mg/kg bw/d, as discussed in Section 8.1 above. Hence, that value should be used to represent the upper bound of the range.

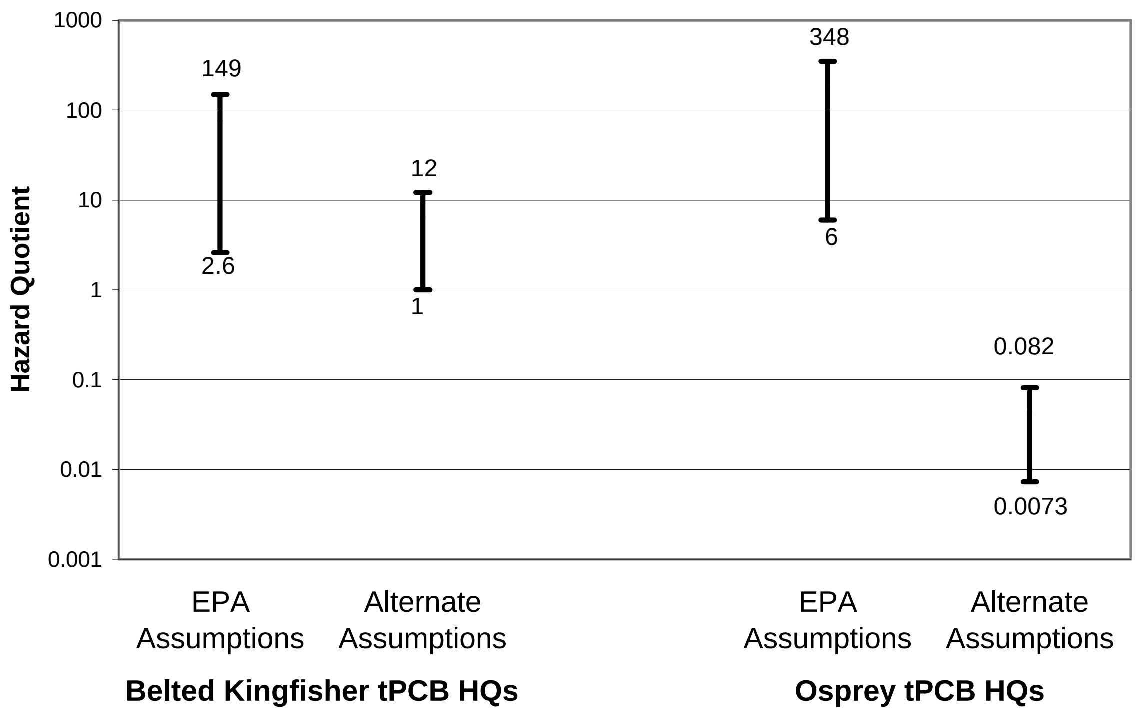
For the lower bound of the range, it is inappropriate to use the dated Lillie et al. (1974) study on white leghorn chickens, which are domesticated and are substantially more sensitive than wild species to PCBs (Bosveld and Van den Berg 1994). Chickens were first domesticated 5,000 years ago in India (Wiley 1997); as such, it is likely that their gene pool has been influenced by selective breeding such that their responses to chemical stressors are not typical of wild species. As recognized in EPA's (1995) *Great Lakes Water Quality Initiative Technical Support Document for Wildlife Criteria*: "many traditional laboratory species…are bred from a fairly homogeneous gene-pool. Use of a [test dose] derived from a 'wildlife' species is thought to provide a more realistic representation of the dose-response relationship which may occur in the natural environment" (EPA 1995, p. 11). The acceptability criteria used to select studies for consideration in the effects assessments (Vol. 2, p. 6-22) should include consideration of domestic/wild status, so that effects metrics based on domesticated species are used only in the absence of any suitable studies on wild species. In this case, a study by Custer and Heinz (1980) on mallards, which was not included in the ERA, provides a suitable study on a wild species. That study yielded a NOAEL of 1.4 mg/kg bw/d tPCBs, which should be used for the lower bound of the range.

Finally, as discussed above for tree swallows, the selection of effects metrics for TEQs inappropriately includes a study based on weekly intraperitoneal injections (Nosek et al. 1992). The only available study that provides a defensible TEQ effects metric for birds is the Hoffman et al. (1996) study, which yielded an effects metric of 25,000 ng/kg bw/d). That value should be used as the only avian effects metric for TEQs.

9.3.3 Correction of Hazard Quotients

Correction of the overly conservative assumptions described above would yield substantially lower HQs for both belted kingfishers and ospreys. For belted kingfishers, the recommended changes to the food intake rate and the tPCB effects metrics would result in up to a 13-fold reduction in the tPCB HQs. For example, for the median tPCB dose for belted kingfishers in Reach 5 (the reach with maximum exposure), these changes would reduce the maximum tPCB HQ from 149 to 12. For ospreys, the recommended changes to the food intake rate, the foraging time assumption (i.e., using an value of 0.5% to reflect the ospreys' migratory status in western Massachusetts), and the tPCB effects metrics would result in reduction of more than three orders of magnitude in the tPCB HQs. For example, for the median tPCB dose to ospreys in the PSA, these changes would reduce the maximum HQ from 348 to 0.082. The impact of these recommended changes for the belted kingfisher and osprey tPCB HQs is illustrated in Figure 9-3 below. Even greater changes would occur in the HQs for TEQs.

Figure 9-3. Impact of Alternative Assumptions on tPCB Hazard Quotients for Piscivorous Birds



Note: Values shown are the upper and lower bound HQs

9.4 Overall Assessment

As noted previously, the ERA concludes that tPCBs and TEQs pose low risks to belted kingfishers in the PSA, because, although the HQs indicate high risk for tPCBs and intermediate risk for TEQs, the field study indicates that these birds are reproducing in the PSA (Vol. 2, pp. 8-48, 8-49; Vol. 6, p. H-63). The

ERA characterizes this conclusion as uncertain because the two lines of evidence do not give concordant results. The ERA also concludes that ospreys are at high risk from exposure to tPCBs and intermediate risk from exposure to TEQs in the PSA, but it characterizes that conclusion as uncertain because only one line of evidence (HQs) is available.

GE concurs with the overall weight assigned to the belted kingfisher field study (Moderate/High). Further, we agree that that study, which showed no evidence of harm, should receive greater weight than the HQs based on modeled exposures and effects. We also believe that use of more defensible food intake rate and effects metrics, as recommended above, would improve the HQ analysis for belted kingfishers, reducing the magnitude of the HQs by approximately 13-fold. Hence, the conclusions of the two lines of evidence for belted kingfishers, while still discordant, would be more closely aligned, thus increasing the certainty in the conclusion of negligible or low risks to these species.

GE also believes that the risks predicted for breeding ospreys that derive all of their prey from the PSA are completely unrealistic, given the absence of any records of breeding osprey in western Massachusetts and the absence of effects observed in a more highly exposed piscivorous bird species (i.e., belted kingfishers). Moreover, the HQs for ospreys rely on unnecessarily conservative exposure assumptions (relating to both food intake rate and foraging distance) and effects metrics that together result in overstatement of risks by more than three orders of magnitude. Additionally, the results from the HQs for ospreys are overextended to speculate that ospreys are not breeding within the PSA due to the presence of contaminants (Vol. 2, p. 8-49) when, in fact, the available records for ospreys demonstrate that western Massachusetts is outside the normal breeding range for ospreys.

As detailed above, breeding ospreys should not be included as a representative receptor of piscivorous birds. As the most highly exposed piscivorous bird species known to breed within the PSA, belted kingfishers adequately represent this feeding guild. However, if it is critical to include a second representative of the feeding guild, great blue herons should be evaluated. In contrast to ospreys, great blue herons are known to breed within two miles of the PSA (which is well within foraging range for this species); and data are available on the productivity of great blue herons breeding within foraging distance of the river relative to those breeding elsewhere in the State, which could provide a second line of evidence to an appropriate HQ analysis.

In any event, even if ospreys are included, the HQs for this species should be recalculated for migrating ospreys, using more realistic exposure and effects assumptions. Such recalculated HQs would be less than 1, indicating negligible risks to migrating ospreys in the PSA.

Based on these lines of evidence, coupled with correction of the deficiencies in the HQ approaches, a more balanced assessment would conclude that piscivorous birds are likely at negligible or low risk from PCBs and TEQs in the PSA.

SUMMARY OF CONCLUSIONS
<u>RISKS TO PISCIVOROUS BIRDS</u>

Based on field data showing no effects on belted kingfishers or great blue herons, and considering the over-conservatism in the HQ analyses, piscivorous bird populations are likely at negligible or low risk from PCBs and TEQs in the PSA. Although ospreys should not be used as a representative species, correction of the HQs for this species would lead to a similar conclusion.

See Table 9-1 for specific comments on Charge Questions 3.5(a)-(j)

Table 9-1. Assessment Endpoint: Survival, Growth, and Reproduction of Piscivorous Birds

EPA Charge Question #3.9	GE Response
(a) Were the EPA studies and analyses performed (e.g., field study, site-specific toxicity study, comparison of exposure and effects) appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), and based on accepted scientific practices?	**Kingfisher HQs:** Basic methodology is appropriate, but overly conservative assumptions are used for both exposure and effects (see below). **Osprey HQs:** Ospreys should not have been selected as a representative species because western Massachusetts is not within the breeding range of ospreys. If a second representative species is required, great blue herons would be a better choice because they do breed within foraging distance of the Housatonic and field data on the productivity of such herons (vs. those breeding elsewhere in the State) are available. (Section 9.2). In addition, the HQs for ospreys use overly conservative assumptions for both exposure and effects (see below).
(b) Were the GE studies and analyses performed outside of the framework of the ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), based on accepted scientific practices, and incorporated appropriately in the ERA?	**Belted Kingfisher Field Study:** This study was appropriate for determining whether the local population of kingfishers is breeding successfully and has a density consistent with the habitat. The ERA advances several criticisms of this study which are unwarranted. Nonetheless, the overall conclusions of the study are appropriately incorporated into the ERA. (Section 9.1)
(c) Were the estimates of exposure appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?	**Belted Kingfisher Field Study:** Uncertainty in the estimates of exposure is overstated. (Section 9.1) **Kingfisher and Osprey HQs:** Food intake rates should be based on published measured rates for free-living belted kingfishers and ospreys, rather than allometric calculations. In addition, the assumption that ospreys derive 100% of their prey from the PSA overestimates exposure by ~200 times, since they are only likely to forage at the PSA during migration, for 1 to 3 days per year. (Section 9.3.1)
(d) Were the effects metrics that were identified and used appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable)?	**Belted Kingfisher Field Study:** Effects information used is appropriate, given purpose of study. (Section 9.1) **Kingfisher and Osprey HQs:** The effects metrics are unsupported. For tPCBs, the upper-bound effects metric should be based on the site-specific tree swallow study, rather than a study of American kestrels, and the lower-bound effects metric should be based on a study on mallards, rather than a study of domesticated chickens. For TEQs, a single effects metric based on Hoffman et al. (1996) should be used, because no other studies are available that offer defensible TEQ effects metrics for birds. (Section 9.3.2)
(e) Were the statistical techniques used clearly described, appropriate (i.e., objective, consistent, and reasonable), and properly applied for the objectives of the analysis?	**Belted Kingfisher Field Study:** Statistical analyses presented in the original report on this study (Attachment L) are appropriate. The ERA's reanalysis of the kingfisher productivity data is flawed (see Attachment A). **HQs:** Mathematical calculations appear appropriate.
(f) Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable)?	**Belted Kingfisher Field Study:** Overall risk characterization is appropriate. **HQs:** Risks from tPCBs are overestimated by 13 times for belted kingfishers and three orders of magnitude for ospreys due to use of inappropriate exposure and effects assumptions. Risks from TEQs are even more substantially overestimated due to use of an unsupported effects metric. (Section 9.3.3).

Table 9-1. Assessment Endpoint: Survival, Growth, and Reproduction of Piscivorous Birds

EPA Charge Question #3.9	GE Response
(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements could be made.	**Belted Kingfisher Field Study:** Uncertainties in this study are overstated. (Section 9.1) **HQs:** Impacts of overly conservative assumptions not adequately recognized. Improvements are recommended in Section 9.3.
(h) Was the weight of evidence analysis appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable)? If not, how could it be improved?	**Belted Kingfisher Field Study:** Weight given to this study (Moderate/High) is appropriate. **HQs:** Due to uncertainties and over-conservatism in assumptions, weight given to the HQs (Moderate) is too high. (Section 9.3)
(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?	Not applicable.
(j) In the Panel members' opinion, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?	**The ERA overstates risks to piscivorous birds in the PSA,** due to its unsupportable selection of breeding ospreys as a representative species (when ospreys do not breed in western Massachusetts) and due to the use of unnecessarily over-conservative assumptions in the HQs. Based on field data showing no PCB effects on belted kingfishers or great blue herons, and considering the over-conservatism in the HQ analyses, piscivorous bird populations are likely at negligible or low risk from PCBs and TEQs in the PSA. Although ospreys should not be used as a representative species, correction of the osprey HQs to apply to migrating ospreys would lead to a similar conclusion. (Section 9.4)

10. PISCIVOROUS MAMMALS (Question 3.7)

Key Points

- The ERA concludes that there are intermediate to high risks to mink and river otter due to PCBs and TEQs in the PSA, as well as downstream of the PSA. This conclusion is based on a number of data interpretations and assumptions that are not well supported by the evidence.

- EPA's mink feeding study does not provide clear evidence of harm due to PCB exposure.

  - Ø It did not show PCB-related effects on many of the endpoints evaluated (e.g., adult food intake rate, breeding success, whelping success, litter size, organ histopathology).

  - Ø Although this study reported an effect on kit survival at 6 weeks at the highest dose level (3.7 mg/kg in fish diet), kit survival at that point was highly variable among all dose groups.

  - Ø Kit mortality prior to 6 weeks cannot be attributed to PCB exposure since no necropsy data are reported for those kits, and necropsies on kits that died later showed that their deaths were due to non-PCB-related causes.

  - Ø GE's statistical reanalysis of the data showed no significant effect on kit survival at the 3.7 mg/kg dose level. Thus, this dose level should be considered an unbounded NOAEL, rather than a LOAEL.

  - Ø In any event, this study demonstrates that mink are less sensitive to the mixture of PCBs present in Housatonic River fish than they are to PCB mixtures present at other sites.

- Field surveys reveal that there is substantial utilization of the PSA by mink and river otter.

  - Ø The ERA's claim that the EPA field survey showed less frequent observations of mink and otter in the PSA than in reference areas or than would be expected is not supported because that survey was too sporadic, did not account for habitat differences, and involved no statistical analysis.

  - Ø The GE survey indicates, based on the spatial and temporal distribution of mink and otter tracks and other signs, that approximately 6-10 mink in 2002, 4-7 mink in 2003, and 2-3 river otter in 2003 (otter were not surveyed in 2002) used the study area regularly or as part of their home range. The ERA's criticisms of that survey are unwarranted.

- The ERA's HQ analyses are inappropriately based on literature-derived effects curves or thresholds, since site-specific effects data are available. Moreover, due to a number of errors in the literature-based effects curves and thresholds, the resulting HQs are overly conservative, as illustrated by the fact that they also predict risks for reference areas.

- **In summary, the existing site-specific data do not demonstrate harm to mink and otter. Based on the literature, there may be risks to mink and otter in the PSA due to consumption of PCB-containing fish at some exposure level. But the ERA understates the threshold and overstates both the magnitude of the risks and the certainty of the conclusions.**

- For reaches downstream of the PSA, the ERA should replace the current MATC of 2.65 mg/kg in fish with a MATC of > 3.7 mg/kg in fish, based on the unbounded NOAEL from the mink feeding study.

10. PISCIVOROUS MAMMALS (Question 3.7)

The ERA employs three types of measurement endpoints to assess risks to piscivorous mammals: (1) a laboratory mink feeding study conducted by EPA contractors using site-specific fish in the feed; (2) field surveys of mink and river otters conducted by both EPA and GE contractors; and (3) risk curves based on modeled exposure and literature-based effects thresholds (i.e., HQs). A MATC derived from the mink feeding study is used to assess risks downstream of the PSA. Based on these lines of evidence, the ERA concludes that mink and otter are at intermediate to high risk from exposure to PCBs and TEQs in the PSA and downstream to Reach 10 for mink and Reach 12 for otter (Vol. 2, pp. 9-78 - 9-80; Vol. 6, pp. I-84, I-85). GE has substantial concerns with the ERA's interpretations of both the mink feeding study and the mink/otter surveys, as well as with the HQ analyses. These concerns are described below, while specific comments related to Questions 3.7(a)-(j) are provided in Table 10-1 (at the end of this section).

10.1 Mink Feeding Study

EPA's mink feeding study was designed to test the hypothesis that exposure to contaminated prey in the PSA may cause adverse effects on the survival, reproduction, and/or growth of exposed individuals (Bursian et al. 2003).[21] Farm-raised adult females were fed a diet containing fish from the PSA at five nominal treatment levels, ranging from 0.25 mg/kg to 4 mg/kg tPCBs (actual concentrations ranged from 0.34 mg/kg to 3.7 mg/kg tPCBs), for two months prior to mating and through mating and whelping of the kits (approximately 160 days). A subset of kits was also fed this diet for six months after whelping. Numerous endpoints were evaluated in this study. These included:

- Adult food intake rate
- Breeding success
- Gestation length
- Whelping success
- Adult body weight
- Litter size
- Organ weight
- Histopathology
- Adult survival
- Kit survival
- Kit body weight
- CYP2B-related activity (BROD and PROD)
- CYP1A1-related hepatic enzyme activities

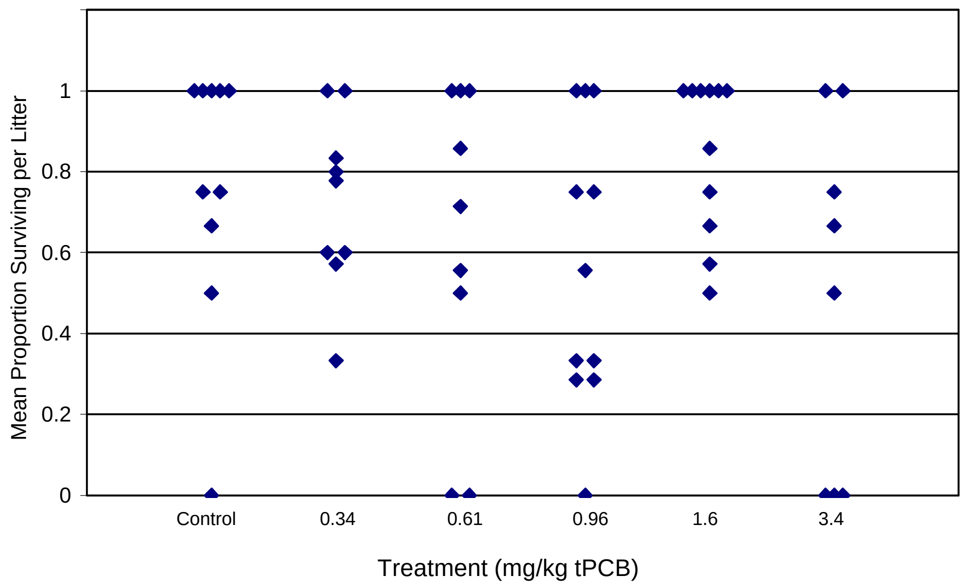
[21] The ERA cites Bursian et al. (2002), which was a draft of the contractors' report on this reproduction/survival study. However, the report contained on the compact disk accompanying the ERA is the contractors' final report dated July 3, 2003, which contains some changes from the 2002 draft. The discussion in the ERA and in these Comments relate to the final report on the reproduction/survival study (Bursian et al. 2003).

No significant PCB-related effects were found for the great majority of these endpoints (Vol. 2, pp. 9-41, 9-42; Vol. 6, p. I-42). Only four statistically significant effects were reported in this study to be associated with tPCBs or TEQs: decreased kit survivability at six weeks, decreased kit body weight at three weeks, and increased induction of CYP1A1-related hepatic enzyme activities (i.e., ethoxycoumarin-O-deethylase or ECOD and ethoxyresorufin-O-deethylase or EROD) (Vol. 2, pp. 9-41, 9-42; Vol. 6, p. I-42, I-43). Effects on weight were transient (i.e., only at one time point) and increased ECOD/EROD induction reflects potential exposure to PCBs, but not adverse effects from that exposure. Hence, these comments focus on the evidence relating to kit survival. As discussed below, our review indicates that survival, in fact, was not adversely affected by PCB exposure.

The ERA reports that kit survival at six weeks of age was significantly reduced at the highest dose (3.7 mg tPCB/kg feed) compared to the control and the second highest dose (1.6 mg tPCB/kg feed) (Vol. 6, Table I.3-2). However, as discussed below, evidence that the dose of 3.7 mg tPCB/kg actually resulted in increased kit mortality is inconclusive due to: (1) the high variability among litters within this treatment group; (2) the lack of necropsies to verify the cause of death in kits up to six weeks old (particularly when contrasted to necropsy data on adult minks and kits that died later, verifying non-PCB-related causes of death); and (3) the fact that there were no statistically significant differences in mortality over the range of exposures evaluated.

Kit survival at six weeks was highly variable among all litters representing each dose groups, as shown on Figure 10-1 below. As an example, in the highest dose group (3.7 mg tPCB/kg), two of the eight litters had 100% survival, one litter each had 50%, 67% and 75% survival, and three litters had 0% survival. Similar variability was seen at lower doses, including the control. Without data at higher doses, it is not clear that mortality at the highest dose was not simply more of this non-PCB-related variability. Moreover, although kit survival was not explicitly (i.e., statistically) evaluated in this study beyond six weeks of age, a subset of 12 randomly selected mink kits were maintained on their respective treatment diets and studied through six months of age. Only two kits died over that period, and both deaths were from causes unrelated to PCBs (Bursian et al. 2003 p. 16 and Table 28).

Figure 10-1. Mink Kit Survival at Six Weeks



Note: x-axis not to scale.
There is no significant effect upon kit survivability at six weeks due to dietary treatment ($F_{(5,132)}$ = 1.49, p = 0.21).

In order to confirm the cause of death, necropsies should have been conducted on all adults and kits that died during the course of the study. However, no necropsy data are presented in either the ERA or the underlying study for kits that died prior to six weeks of age. By contrast, necropsies were conducted on all adult mink that died during the course of the study, as well as kits that died after six weeks of age. Those necropsies indicated that the mortality was not due to PCB toxicity (Bursian et al. 2003, Tables 27 and 28). Instead, it was attributed to infections common in captive mink – urolithiasis and hemorrhagic cystitis. Given the presence of these infections in both adult mink and mink kits that died after six weeks of age, the natural conclusion is that deaths in kits less than six weeks of age were likewise attributable to causes other than PCBs. Without evidence to support a contrary conclusion, it is inappropriate to attribute mortality in kits that died before six weeks of age to PCBs.

Insufficient detail regarding the results of the statistical analysis are provided in the ERA (Vol. 2, p. 9-41; Vol. 6, p. I-42, Table I.3-2) and the underlying report (Bursian et al. 2003) to allow verification of the reported statistically significant reduction in kit survival at the highest dose at six weeks. In an attempt to reproduce this result, GE conducted an independent ANOVA of this data, testing the relationship between total live kits at six weeks relative to the total number of kits whelped and the PCB treatment. The results of this analysis are presented in detail in Attachment N to these Comments and are shown at the bottom of Figure 10-1 (above). This analysis indicates there was no significant effect upon kit survivability due to dietary treatment at six weeks. As a result, 3.7 mg tPCB/kg feed (0.414 mg/kg bw/d) represents a NOAEL and not a LOAEL, as reported in the ERA (Vol. 2, p. 9-78; Vol. 6, p. I-47).

Other studies have reported reproductive failure or mortality in offspring (e.g., Aulerich et al. 1985; Heaton et al. 1995; Hornshaw et al. 1983; Platanow and Karstad 1973) at concentrations lower than the highest treatment dose in this study (on a tPCB basis). As discussed in the ERA, these data suggest that the congener mix in Housatonic River fish is not as toxic as those used in other studies, or that other COPCs also present in feed used in other studies may have contributed to the effects observed in those studies (Vol. 2, pp. 9-76, 9-77; Vol. 6, p. I-47).

ASSESSMENT OF EPA'S MINK FEEDING STUDY

- **This study does not provide evidence of a dose-response relationship between decreased kit survival and tPCBs.**
- **The kit mortality that was observed cannot be attributed to exposure to tPCBs because no necropsies were conducted to verify the cause of death, and deaths of older kits and adults were found to be attributable to causes other than PCBs.**
- **A statistical reanalysis of the data indicates an unbounded NOAEL of 3.7 mg/kg tPCBs for kit survival.**
- **Overall, this study demonstrates that mink are less sensitive to the mixture of PCBs present in Housatonic River fish than they are to PCB mixtures present at other sites.**

10.2 Jaw Lesion Study

In a separate report, Bursian and Yamini (2003) reported the results of an evaluation of the histopathology of jaws of 36 six-month-old kits from the feeding study. Skulls of these kits were examined to determine if the dietary treatments induced jaw lesions that had been observed previously by other investigators (Render 2001, 2002) in the same Michigan State University laboratory where the Bursian et al. study was conducted. Although the same dietary treatments were used in the jaw lesion and the mink feeding study, Bursian and Yamini (2003) reported the dose for this study as PCB 126 (instead of tPCBs or TEQs) to facilitate comparisons with other studies that also evaluated this effect. The ERA reports that the three highest PCB 126 doses evaluated in this study (0.054 ng/g, 0.098 ng/g, and 0.41 ng/g) led to the proliferation of maxillary and mandibular squamous cells in mink, but no measurable effects in the form of gross abnormalities (Vol. 2, pp. 9-42, 9-43; Vol. 6, p. I-44). Nonetheless, the ERA suggests that exposure to PCB-126 over a longer period of time would increase the severity of these lesions and lead to the loss of teeth, and ultimately could lead to starvation (Vol. 2, pp. 9-43, 9-79; Vol. 6, pp. I-44, I-85). As discussed below, these conclusions overstate the data in this study and in the literature.

Bursian and Yamini (2003) also reported the results of another unpublished study in which mink kits exposed to 0.24 ng/g PCB 126 in feed in the same manner as used in this study also had a proliferation of periodontal squamous epithelial cells, but no gross abnormalities. The only reported incidence of gross abnormalities in mink kits exposed to PCBs was at a much higher concentration than used in the current study. Render et al. (2001) reported loose and displaced incisor teeth in kits exposed to 24 ng/g PCB 126 in feed. This dose was 58.5 times higher than the highest dose used by Bursian and Yamini (2003) (i.e., 0.41 ng/g PCB-126). It is inappropriate to extrapolate effects documented at this substantially higher dose to doses used in the Bursian and Yamini (2003) study. While this endpoint may merit additional research, the Bursian and Yamini (2003) site-specific study does not define an adverse effects threshold. Moreover, the ERA's speculation that the effects observed (i.e., proliferation of maxillary and mandibular squamous cells) could ultimately cause the animals to die of starvation has not been documented by these studies (Bursian and Yamini 2003; Render et al. 2001, 2002). There was no evidence of wasting or other related problems in kits followed to six months of age in the mink feeding study (Bursian et al. 2003; Bursian and Yamini 2003).

10.3 EPA and GE Mink and River Otter Surveys

Both EPA and GE contractors conducted mink and river otter surveys over a span of several years in the PSA. The qualitative nature of these surveys limits conclusions that can be drawn regarding the status of local populations. Nonetheless, these surveys generally indicate that mink and river otter are present in the PSA (Vol. 2, pp. 9-49 - 9-52; Vol. 6, pp. I-55 - I-63) and the GE study concluded that both mink and river otter use the PSA as part of their home ranges (Bernstein et al. 2003). The ERA concludes that the EPA survey indicates evidence of harm with a high magnitude of effect, that the GE survey indicates no evidence of harm with a low magnitude of effect, and that the EPA survey is entitled to more weight than the GE survey (Vol. 2, Tables 9.5-4, 9.5-5; Vol. 6, Tables I.4-5, I.4-6). GE believes that the ERA's conclusions represent an incorrect interpretation of the survey data, as discussed below. These points are supported by Attachment O, which is a full report on GE's survey, and Attachment P, which contains comments by Professor Michael Chamberlain (an expert on carnivorous and piscivorous mammals and a co-author of the report on GE's survey) on the ERA's interpretation of both the EPA and GE surveys.

10.3.1 EPA mink and otter survey

EPA contractors conducted surveys in the PSA and in several reference areas in the winters of 1998-1999 (termed 1999 hereafter) and 1999-2000 (termed 2000 hereafter). Six transects were established in the PSA for snow tracking in 1999 and three transects were established in 2000. Transects were also established in two reference areas in 1999 and four reference areas in 2000. These transects were surveyed after a fresh snowfall for a minimum of two or three snow events each winter (Vol. 6, p. I-56). Three scent post station transects were also monitored on December 15-17, 1998, in late summer and early fall 1999, and during winter snow tracking in 2000. The EPA survey concluded that mink and river otter were observed much less frequently in the PSA than in the reference areas or than would be "expected" (Vol. 2, pp. 9-50 - 9-51; Vol. 6, p. I-60). For a number of reasons, discussed below and by Dr. Chamberlain (in Attachment P), this survey does not provide reliable evidence of harm:

- Although the ERA characterizes this survey as being conducted over a four-year (1998-2001) period (Vol. 2, p. 9-49), it was in fact only conducted on sporadic occasions during two winters, with scent post transect monitoring on a few other occasions, as described above. No data were collected in 2001.

- No statistical analyses (e.g., chi-square) were conducted to support the comparison to reference sites, and no "expected" values (i.e., frequency of observations of mink and otter sign) are presented.

- The comparisons do not account for habitat differences between the PSA and the reference areas (lakes and ponds that are unlike the complex riverine system of the PSA), which are known to influence mammalian movements, space use, and other population parameters.

- Factors that could have affected the survey success between locations and survey periods were not accounted for (e.g., differences in access due to habitat, skill of personnel, randomization of sampling design).

- The only documentation provided for mink tracks reported in the ERA (Vol. 3, p. 6-12) is one picture of mink tracks (Vol. 3, Figure 6-6).

10.3.2 GE mink and otter survey

The objectives of the GE survey (Bernstein et al. 2003) were to qualitatively determine the presence/absence, abundance, and distribution of free-ranging piscivorous mammals between New Lenox Road and Woods Pond (the Survey Area), including areas adjacent to the mainstem of the river and nearby tributaries. Mink were the focus of the survey between spring 2001 and spring 2002. The primary methods used to determine the presence of mink were the monitoring of tracks at scent post stations in the summer and fall (2001 only) and in the snow during the winters of 2001-02 (termed 2002 hereafter) and 2002-03 (termed 2003 hereafter). The survey was expanded to include river otter in the winter of 2003, when river otter tracks and slides were also monitored. Attachment O presents: (1) the full survey report; (2) the resumes of the investigators; (3) maps indicating the location of all tracks observed during the survey; and (4) photographs of mink and otter taken using motion-sensing cameras (e.g., Figures 10-2 and 10-3 below), as well as photographs of the tracks observed.

Mink signs were consistently observed in the Survey Area during the winters of 2002 and 2003. (Figure 10-4 below illustrates, for example, the locations of the mink tracks identified in the winter of 2003.) The spatial and temporal patterns of mink tracks indicated that approximately 6-10 mink in 2002 and approximately 4-7 mink in 2003 likely used the Survey Area as part of their home range. In 2003, 41 river otter tracks were observed in the Survey Area, along with six sprainting stations (i.e., latrines) and two den sites (one confirmed, one potential). Multiple photographs of otter were taken near the den sites. Based on this evidence, the study concluded that two river otters used the Survey Area as their core use area in 2003, that another one used the Survey Area regularly, and that at least one other otter passed through the Survey Area.

Figure 10-2. Mink Photographed on the Mainstem Using Motion Detector Camera in Winter 2003

Figure 10-3. River Otter Photographed on the Mainstem Using Motion Detector Camera in Winter 2003

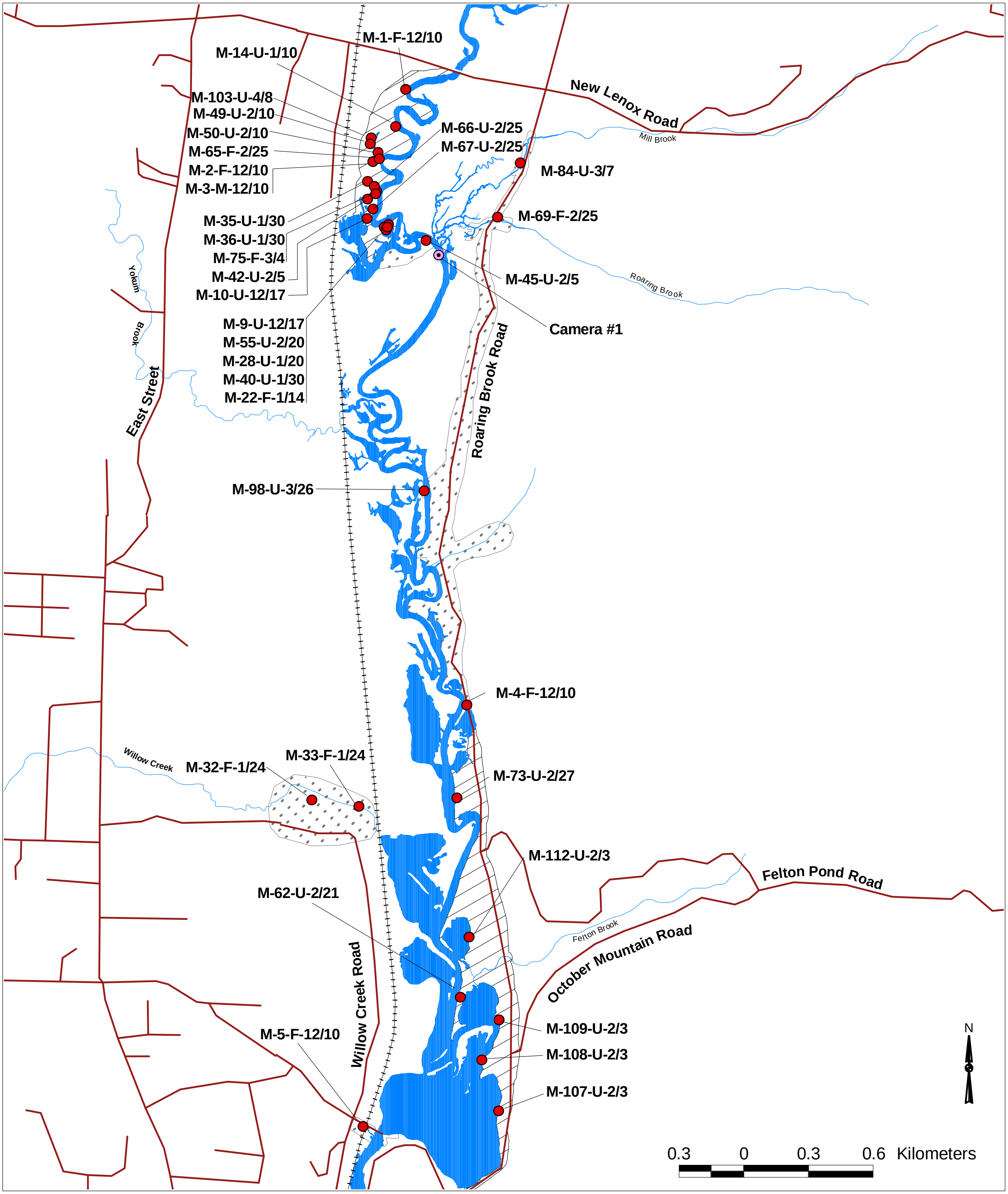


Figure 10-4

Mink Track Locations (Winter 2003)

Housatonic River, Massachusetts

Legend

Mink track location

Camera location

Labels

M-xx: Mink
-M: M = male, F = female,
U = unknown
-1/14: month/day (e.g. Jan. 14)

Searching transects (frequent)

Searching transects (infrequent)

The ERA advances several criticisms of this survey, which it claims undermine the conclusions (Vol. 2, pp. 9-51, 9-52; Vol. 6, pp. I-61 - I-63). These criticisms are unwarranted. The more significant criticisms are discussed below and by Dr. Chamberlain in Attachment P.

- The ERA notes that the great majority of the mink and otter signs were observed in winter, and it argues that this finding suggests that these mammals are not present in the Survey Area in other seasons. In fact, however, more mink tracks were documented in the months with snow because fresh snow cover provided an extensive area over which animals could be more easily tracked than in snow-free months, when monitoring was limited to artificially created tracking surfaces. For that reason, less effort was spent surveying in the snow-free months (i.e., scent stations were only used in summer/fall 2001; motion-tracking cameras were deployed more than three times more often in winter months than in snow-free months). Moreover, scent stations are generally less successful for these mammals in the summer due to dispersal patterns and are generally less effective for secretive piscivores like mink than for more curious generalists (see Attachment P). Finally, otter signs were only monitored in the winter; therefore, it is inappropriate to comment on the presence or absence of otter signs in the snow-free months.

- The ERA also states that the mink and otters observed in the Survey Area were likely transients. There is no basis or support for that conjecture. Based on the temporal and spatial distribution of the tracks, as well as other mink and otter signs, the GE survey provided estimates of the number of mink and river otters that have home ranges that *include* the PSA. These animals cannot be confirmed as residents or simply users of the area, and it is thus inappropriate to assume that they were transients.

- The ERA claims that the field personnel "lack[ed] tracking expertise and experience" (Vol. 2, p. 9-52). There is no basis for this slight. The primary expert in track identification retained by GE, Mr. Bernstein, is a former New York Conservation Officer with decades of experience in identifying wildlife tracks, as shown by his resume (Attachment O, Exhibit A). Further, Dr. Chamberlain found him and the other field biologists to be fully competent (Attachment P). The ERA's claim is wholly unwarranted.

- The ERA criticizes the survey for lack of reference sites. However, the objective of the study was simply to document the presence of piscivorous mammals in the Survey Area. The results were put into the context of mink and otter densities reported in the literature.

- The ERA criticizes the level of track documentation. However, more than half of the tracks observed in 2002 and all tracks reported in 2003 were photographed and their locations were documented using global positioning system (GPS) (see Attachment O, Exhibits O-2 through O-4);

- The ERA questions the survey findings on the ground that "[n]o confirmed tracks were observed in the presence of EPA oversight personnel" (Vol. 6, p. I-63). This is not surprising as EPA observers were rarely in the field with GE field biologists (e.g., only 3 of 24 days in the winter of 2003). Given Mr. Bernstein's experience and the ample photo documentation, EPA's absence in the field when the tracks were found is immaterial to the quality of the study.

In the weight-of-evidence evaluation, the ERA assigns the GE survey lower overall weight (Moderate) than the EPA survey (Moderate/High), as well as lower ratings for many individual attributes (Vol. 2, Table 9.5-3; Vol. 6, p. I-70 - I-78). Limitations in study methodology, interpretation, and data quality are frequently cited as reasons for these lower weights. However, as discussed previously, the ERA misinterprets the GE survey, and in some cases (e.g., for temporal representativeness) EPA's own survey, resulting in a lower weight for the GE survey. GE believes that a more balanced weight-of-evidence evaluation would indicate that the more extensive GE survey warrants greater weight than the EPA survey. However, at a minimum, it certainly does not warrant lower weight.

10.4 Modeled Exposure and Effects

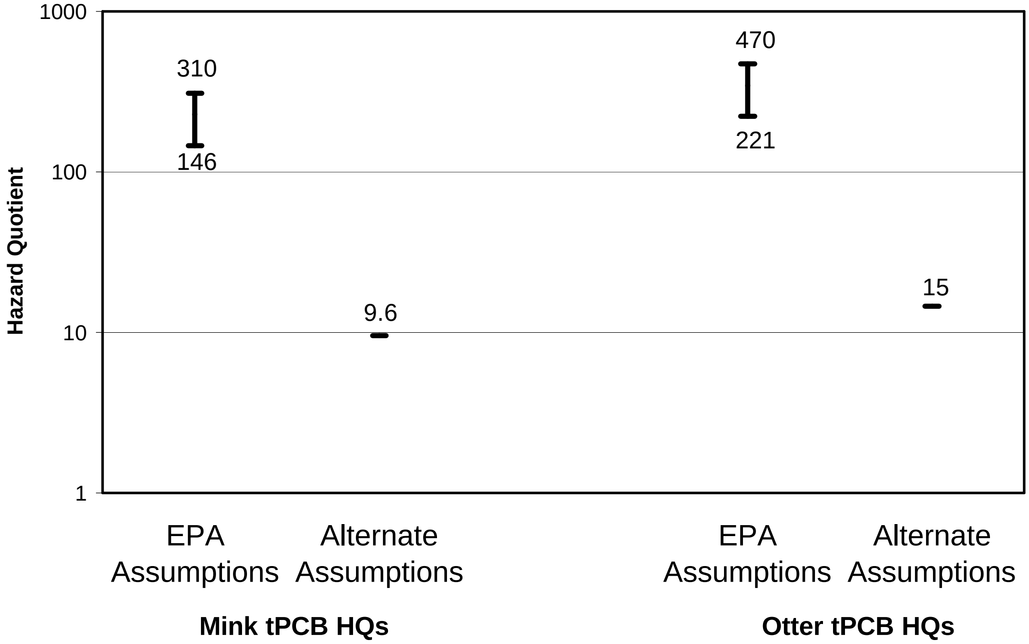
Risks to mink and otters in the PSA are also assessed using modeled exposure data together with a dose-response curve for tPCBs and an effects threshold range for TEQs, both derived from literature-based studies that are not site-specific (i.e., HQs) (Vol. 2, pp. 9-47, 9-48; Vol. 6, p. I-51 - I-54). Given the availability of a site-specific, species-specific study (Bursian et al. 2003), literature-based effects information should not have been used to model risk. The use of non-site-specific studies is of particular concern given that the EPA mink feeding study, as noted above, demonstrates that the congener mix in the Housatonic River fish is less toxic to mink than those tested in other feeding studies, thus indicating that the effects data from those other studies have limited, if any, applicability to this site.

In any event, there are several problems with the literature-based effects curve and thresholds. The tPCB dose-response curve is based on studies by Bleavins et al. (1980) and Aulerich et al. (1985) (Vol. 1, pp. 9-47, 9-48). However, the doses reported in that curve may be incorrect. Because the majority of the dose-response data used to derive effects thresholds are presented in the literature by dietary dose (e.g., mg/kg tPCBs in feed), it is necessary to convert the reported doses by normalizing them to 1 kg body weight (i.e., mg/kg bw/d). The assumptions used in these normalizations are not provided in the ERA. We attempted to re-create the doses reported in the tPCB risk curve based on body weight and food intake rate assumptions provided in the studies cited. The normalized doses resulting from this independent analysis are not consistent with those reported in the dose-effect graph for tPCBs shown in the ERA (Vol. 2, p. 9-48). For instance, the lowest concentration of tPCBs in the diet that was used by Bleavins et al. (1980) is 5 mg/kg. Assuming a food ingestion rate of 150 g/d and a body weight of 800 g (Bleavins et al. 1980, p. 632), the resulting dose is 0.9375 mg/kg bw/d. This dose is more than three times higher than the lowest dose (0.3 mg/kg bw/day) on the risk curve associated with the effect (no kits per female) observed in all of Bleavins et al.'s (1980) tPCB treatments. An additional problem with the tPCB risk curve is that, based on comparison of the number of data points on the curve with the data from the underlying studies, the curve includes the results of congener-specific treatments from Aulerich et al. (1985), whereas it is supposed to be limited to tPCB effects.

The TEQ effects threshold range is substantially more conservative than the tPCB risk curve because it was based on reduced growth (Vol. 2, p. 9-48; Vol. 6, p. I-53, I-54), which in some cases (e.g., Hochstein et al. 2001) is transitory (i.e., occurred at an early life stage but not later stages) and in any event is not directly linked to reproductive success or population-level effects. In contrast, the tPCB risk curve is based on fecundity (i.e., kits born per female), which is a direct measurement of reproductive success. The study used to support the low TEQ effects threshold of 3.6 ng/kg bw/d based on growth (Heaton et al. 1995) also measured fecundity directly, and the lowest concentration at which statistically significantly adverse effects were observed was 10.7 ng/kg bw/d TEQ, a dose almost three times higher than the growth-based threshold used in the ERA. The threshold reported in that study also likely overestimates risks when applied to mink in the PSA, because it was based on a diet of fish from Saginaw Bay, Michigan, which is contaminated with a number of chemicals other than PCBs.

Due to these errors and inconsistencies, the HQs overestimate risks to piscivorous mammals. This over-conservatism is illustrated by the fact that the HQs also predict risks for reference areas (Vol. 2, Table 9.5-2; Vol. 6, Table I.4-3). As previously discussed, it is not appropriate to use literature-based effects metrics to assess risks to mink and river otter in the PSA when site-specific toxicity data are available. As also shown above, those data would support an unbounded NOAEL of 3.7 mg/kg tPCBs (68.5 ng/kg TEQs) in diet for this site. Using this NOAEL as a conservative basis for deriving effects metrics would yield site-specific dose-based effects metrics of 0.414 mg/kg bw/d for tPCBs and 7.67 ng/kg bw/d for TEQs. Use of these effects metrics in the HQs for mink and otter would result in an approximate 30-fold reduction for the tPCB HQ and it would reduce the TEQ HQ by more than half. For example, based on the median modeled exposure estimate for mink in Reach 5 (3.97 mg/kg bw/d for tPCBs and 147 ng/kg bw/d for TEQs – Vol. 6, Tables I.2-6, I.2-7), use of the site-specific effects metrics would reduce the HQs from 310 to 9.6 for tPCBs and from 40 to 19 for TEQs. Similarly, the HQs for otter would be reduced from approximately 470 to 14 for tPCBs and from approximately 20 to 9 for TEQs. The impact of these differences in effects metrics on the mink and otter HQs for tPCBs is shown on Figure 10-5.

Figure 10-5. Impact of Alternative Assumptions on tPCB Hazard Quotients for Piscivorous Mammals



Note: Values shown are the upper and lower bound HQs

A review of the HQs for the reference locations provides additional evidence that these revised thresholds are more appropriate. For example, the tPCB HQ for mink in Threemile Pond (a reference site) is over 5 using the lower-bound literature-based effects metric; using the site-specific data, this HQ is reduced to 0.17, which is more consistent with the estimate of risk one would expect in a reference area. TEQ HQs would continue to exceed 1.0 for mink at Threemile Pond and otter in the upstream reference, but they would be reduced by half to less than three for each.

10.5 Overall Assessment

The ERA concludes that mink and river otters are at an intermediate to high risk from tPCBs and TEQs in the PSA (Vol. 1, pp. 9-78 - 9-80; Vol. 6, p. I-84). However, key interpretations on which this conclusion is based are not supportable. Based on a statistical reanalysis of the kit survival data from the mink feeding study, that study actually provides undetermined evidence of harm, for both tPCBs and TEQs. In assessing the mink and otter field surveys, the ERA does not adequately take account of the problems in EPA's survey and inappropriately discounts the findings from the GE survey, resulting in an unbalanced weighting of those surveys. The HQs provide evidence of harm, but overestimate the magnitude of response due to their reliance on literature-based effects metrics rather than site-specific data, as well as due to other errors and inconsistencies in interpreting the underlying data.

In short, review of the literature indicates that there may be risks to mink and otter in the PSA due to their consumption of PCB-containing fish above some exposure level. However, based on the site-specific data, the ERA adopts a threshold that is too low for this site and it overstates both the magnitude of the risks and the certainty of the conclusions.

SUMMARY OF CONCLUSIONS
RISKS TO MINK AND OTTER IN PSA

The mink feeding study (as reanalyzed) did not show adverse effects on mink at 3.7 mg/kg PCBs in diet, and the field survey data showed considerable usage of the PSA by minks and otter. Based on the literature, there may be risks to mink and otter in the PSA due to consumption of PCB-containing fish at some exposure level. However, the threshold must be > 3.7 mg/kg in fish, and any risk conclusions are uncertain given the lack of site-specific data showing harm to mink and otter.

10.6 Extrapolation to Downstream Reaches

The ERA estimates risks to piscivorous mammals downstream of Woods Pond by comparing concentrations of tPCBs in prey fish collected from Reaches 7 through 16 to a MATC of 2.65 mg/kg tPCBs in fish (whole body, wet weight) (Vol. 2, p. 9-78; Vol. 6, p. I-83). This MATC is the geometric mean of the two highest doses used in the site-specific feeding study (Bursian et al. 2003), which were erroneously determined to be the NOAEL and LOAEL from the study. As discussed previously (Section 10.1), because no effects were observed at the highest treatment level in this study, the assessment of risk downstream of the PSA should be based on a NOAEL of 3.7 mg/kg tPCBs in fish (i.e., MATC >3.7 mg/kg tPCBs), rather than the MATC of 2.65 mg/kg tPCBs that is currently used.

See Table 10-1 for specific comments on Charge Questions 3.7(a)-(j)

Table 10-1. Assessment Endpoint: Survival, Growth and Reproduction of Piscivorous Mammals

EPA Charge Question #3.7	GE Response
(a) Were the EPA studies and analyses performed (e.g., field study, site-specific toxicity study, comparison of exposure and effects) appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), and based on accepted scientific practices?	**EPA Mink Feeding Study:** Basic design and methodology were appropriate. **EPA Mink/Otter Field Survey:** Basic methodology was appropriate, but comparisons between PSA and reference sites do not account for habitat differences and no statistics support the analysis of the data. (Section 10.3.1) **HQs:** Basic methodology is appropriate, but HQs should not have been based on effects metrics from the literature when site-specific effects data are available. (Section 10.4).
(b) Were the GE studies and analyses performed outside of the framework of the ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), based on accepted scientific practices, and incorporated appropriately in the ERA?	**GE Mink/Otter Field Survey:** This survey was appropriate for qualitatively determining the presence, abundance, and distribution of mink and otter in the PSA. The basic methodology and results of the study are misinterpreted and mischaracterized in the ERA. The weight assigned is too low. (Section 10.3.2)
(c) Were the estimates of exposure appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?	**EPA Mink Feeding Study:** Estimates of exposure are reasonable. **EPA and GE Mink/Otter Field Surveys:** Exposure is not estimated. **Modeled Exposure and Effects:** Estimates of exposure are reasonable.
(d) Were the effects metrics that were identified and used appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable)?	**EPA Mink Feeding Study:** Although the ERA asserts that this study found effects on kit survival at 6 weeks at an exposure level of 3.7 mg/kg tPCBs in diet, an independent review and analysis of the data indicate no reliable evidence of PCB effects on kit survival at that level. Hence, the effects metric based on this study should be an unbounded NOAEL, rather than a LOAEL, of 3.7 mg/kg tPCBs. (Section 10.1) **HQs:** The dose-response curve for tPCBs and the effects threshold range for TEQs were inappropriately derived from literature-based studies, rather than site-specific studies. Given the lesser sensitivity of mink to Housatonic River PCBs (as shown by the mink feeding study) compared to that reported in other studies, this results in an overestimate of risk. In addition, there are a number of errors in the literature-based effects curve and thresholds, which tend to overestimate risks. (Section 10.4).
(e) Were the statistical techniques used clearly described, appropriate (i.e., objective, consistent, and reasonable), and properly applied for the objectives of the analysis?	**EPA Mink Feeding Study:** The ERA's finding of a statistically significant decrease in survival in mink kits at 6 weeks at the highest exposure level could not be confirmed by an independent statistical analysis. (Section 10.1). **HQs:** There appear to be errors in the calculation of the dose-response curve for tPCBs. (Section 10.4)

Table 10-1. Assessment Endpoint: Survival, Growth and Reproduction of Piscivorous Mammals

EPA Charge Question #3.7	GE Response
(f) Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable)?	**EPA Mink Feeding Study:** The ERA does not properly characterize risk based on this study, because this study does not provide reliable evidence of adverse effects from PCBs on minks, even at the highest exposure level in the study. (Section 10.1) **EPA Field Survey:** The ERA's claim that this survey showed less frequent observations of mink/otter in the PSA than in reference areas or than would be "expected" is not supported because that survey was too sporadic, did not account for habitat differences between PSA and reference sites, and involved no statistical analysis. (Section 10.3.1) **GE Field Survey:** The ERA correctly concludes that this study showed no evidence of harm, but it advances several criticisms of this study that are unwarranted. (Section 10.3.2) **HQs:** The HQs based on literature thresholds overestimate risks by about 32 times for tPCBs and 2 times for TEQs, compared to HQs based on conservative site-specific effects metrics derived from the mink feeding study. (Section 10.4)
(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements could be made.	**EPA Mink Feeding Study:** Uncertainties associated with high variability in observed survival in the highest dose group at 6 weeks and the lack of corroborating evidence from necropsies to verify whether observed mortality was related to PCBs are not considered. (Section 10.1) **GE Field Survey:** Uncertainties are overstated due to unjustified criticisms of the study. (Section 10.3.2) **HQs:** Uncertainties associated with not using site-specific data, particularly given the apparent lesser sensitivity of mink to PCBs in the PSA relative to other literature studies, are not considered. (Sections 10.1 and 10.4)
(h) Was the weight of evidence analysis appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable)? If not, how could it be improved?	**EPA Mink Feeding Study:** Weight given to this study (High) is appropriate, but the data are not properly interpreted. They actually show undetermined evidence of harm. (Sections 10.1 and 10.5) **EPA and GE Field Surveys:** Weights given to EPA survey (Moderate/High) and GE survey (Moderate) are unbalanced given greater coverage (spatially and temporally) of GE survey. GE survey should have higher weight. (Sections 10.3.2 and 10.5) **HQs:** Weight given to literature-based HQs (Moderate/High) is too high given uncertainties and problems in effects metrics based on literature and because these HQs do not account for differences in the sensitivity of mink to PCBs in the PSA compared to literature studies. (Sections 10.1 and 10.4)
(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?	The ERA overestimates the risks to mink and otter in downstream reaches because it relies on a MATC that is too low. The MATC should be based on an unbounded NOAEL of 3.7 mg/kg tPCBs from the mink feeding study. (Section 10.6)
(j) In the Panel members' opinion, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?	**The ERA's conclusion that piscivorous mammals are at high risk from PCBs in the PSA is not supported by the site-specific data.** Based on the literature, there may be risks to mink and otter in the PSA due to consumptions of PCB-containing fish at some exposure level. However, based on the site-specific study, the threshold must be > 3.7 mg/kg in fish, and any risk conclusions are uncertain given the lack of site-specific data showing harm to mink and otter. (Section 10.5)

11. OMNIVOROUS AND CARNIVOROUS MAMMALS (Question 3.6)

Key Points

Short-Tailed Shrews

- For short-tailed shrews (and other small mammals), EPA's small mammal trapping study showed no evidence of differences in placental scars across sites with a wide range of exposure levels. Although this study has limitations, it provides no evidence of harm, rather than undetermined evidence as stated in the ERA.

- A site-specific short-tailed shrew population demography field study conducted by a small mammal expert on behalf of GE on floodplain grids with varying soil tPCB concentrations showed high densities of short-tailed shrews in the PSA and no relationships between any demographic parameter measured (population density, survival, sex ratio, reproduction, growth, and body mass) and tPCB concentrations in soil in the grids. The study also showed that values for those parameters were within ranges reported in the literature.

  - Ø EPA reanalyzed the data from this study using different tPCB soil concentrations for the grids, and has found a weak, but significant, relationship between those concentrations and survival of the shrews. The ERA does not provide sufficient details regarding this statistical reanalysis.

  - Ø However, a further reanalysis of the data by the study author using the tPCB concentrations estimated by EPA and the same statistical technique used by EPA found no significant relationship between tPCB concentrations and survival.

- The ERA's HQ analyses for shrews used generic modeled food intake rates for mammals when measured species-specific rates are available, and effects metrics based on rodents. As a result, the HQs are highly uncertain and conservative.

- The ERA mischaracterizes the results of the field studies as showing undetermined evidence of harm when they in fact showed no evidence of harm, and it gives too much weight to the HQ analyses. **A more balanced evaluation of the evidence would indicate that shrews are at low, if any, risk from PCBs in the PSA.**

Red Fox

- The EPA field survey, which indicates the presence of red fox in the PSA, does not provide evidence of harm to red fox.

- The HQ analyses for red fox are highly uncertain for the same reasons as apply to the shrew HQs – the use of generic exposure assumptions and surrogate toxicity data for effects metrics.

- The ERA concludes in some places that risks to red fox in the PSA are intermediate but uncertain, and in other places that they are undetermined. GE concurs with the latter conclusion, primarily because **there are insufficient data to draw defensible conclusions about the presence and magnitude of harm to red fox in the PSA**.

11. OMNIVOROUS AND CARNIVOROUS MAMMALS (Question 3.6)

The ERA states that three lines of evidence are available to assess risks to omnivorous and carnivorous mammals: (1) field surveys conducted by EPA consultants; (2) a site-specific population demography study of short-tailed shrews, conducted by an expert retained by GE; and (3) HQs based on modeled exposure and effects for the short-tailed shrew and the red fox. The ERA's conclusions from these lines of evidence are internally inconsistent. For the short-tailed shrew, the ERA states in some places that the evidence shows intermediate (but uncertain) risks from both tPCBs and TEQs in the PSA (e.g., Vol. 2, pp. 10-63, 10-67), and in other places that it shows risks from tPCBs but little or no risks from TEQs (e.g., Vol. 2, pp. 10-59, 10-65). For the red fox, the ERA states in some places that there are intermediate (but uncertain) risks from PCBs and TEQs in the PSA (e.g., Vol. 2, p. 10-63) and in other places that the risks are undetermined (e.g., Vol. 2, pp. 10-59, 10-67). In any case, GE has a number of concerns about the ERA's interpretation and weighting of the various lines of evidence. Those concerns, as well as the conclusions that GE believes should be drawn from the evidence, are discussed below, and a summary of specific comments related to Questions 3.6(a)-(j) are provided in Table 11-2 (at the end of this section).

11.1 EPA Field Surveys

EPA contractors conducted semi-quantitative trapping of small mammals at three locations in September of 1998 and at three additional locations in August and September 1999 (Vol. 6, p. J-44). The objectives of these efforts were to verify the relative abundance of different species in the PSA, to provide tissue samples for PCB analysis, and to determine the reproductive status of females. Short-tailed shrews were the most abundant species captured. Trapped small mammals (including short-tailed shrews) were evaluated for placental scars. As indicated in the ERA, the use of placental scars to assess reproductive success has limitations (Vol. 2, p. 10-42; Vol. 6, p. J-47). Nonetheless, there was no evidence of differences in numbers of placental scars across sites, despite the wide range of exposure (as documented by tissue concentrations ranging from 4 to 148 mg/kg tPCBs). Thus, this aspect of the survey, despite its limitations, actually provides no evidence of harm, rather than undetermined evidence as stated in the ERA (Vol. 2, p. 10-58).

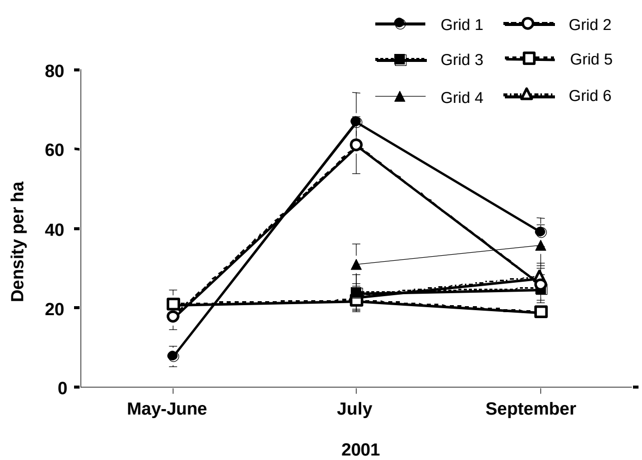
The ERA also notes that both large and small mammals were surveyed over a four-year period between 1998 and 2001 (Vol. 2, pp. 10-41, 10-42; Vol. 6, p. J-44). The ERA reports the presence of red fox throughout the PSA (Vol. 6, p. J-11), but provides no other details regarding these field surveys. The ecological characterization indicates that the red fox was the second most commonly observed large mammal in the PSA (Vol. 3, Section III, p. 6-15). This study was not sufficiently quantitative to provide evidence regarding harm or to be accorded the same weight (Moderate/High) (Vol. 2, p. 10-58; Vol. 6,

Table J.4-7) as the substantially more rigorous GE short-tailed shrew demography study (discussed below).

11.2 GE Population Demography Field Study of Short-tailed Shrews

In 2001, Dr. Rudy Boonstra of the University of Toronto (on behalf of GE) examined the demography of local populations of short-tailed shrews inhabiting the Housatonic River floodplain in areas with varying PCB concentrations in soil, some designated low (1 to 3 mg/kg) and some high (17 to 38 mg/kg) (Boonstra and Bowman 2003).[22] The objective of the study was to assess whether PCBs were adversely affecting the population demography of these small mammals living in their natural environment. Populations were intensively live-trapped on 1 hectare (ha) grids in two or three sampling events conducted between spring and autumn 2001. The study showed no statistical relationships between spatially averaged tPCB sediment concentrations in the grids and any of the demographic parameters measured (population density, survival, sex ratio, reproduction, growth, and body mass). Further, densities were high (usually exceeding 20/ha, and on two grids exceeding 60/ha in summer) (see Figure 11-1 below), survival was good (typically 60 to 75% per 30 days), and sex ratio, reproduction rates, growth rates and body mass were all within ranges reported in the literature. Thus, there was no evidence of adverse effects on these shrew populations as a result of exposure to PCBs. A manuscript describing this study (Boonstra and Bowman 2003), published in the June 2003 issue of *Environmental Toxicology and Chemistry*, is provided as Attachment Q to these Comments.

Figure 11-1. Population Densities (± 1 SE) per ha of *Blarina brevicauda* in 2001 on Six Live-Trapping Grids Along the Housatonic River, Massachusetts Between Pittsfield and Woods Pond



Note: Solid symbols indicate sites with high PCB concentrations and open symbols indicate sites with low PCB concentrations. Points without error bars have very narrow SE values obscured by the point itself.

22 The range of tPCBs in shrew tissue samples collected by EPA in conjunction with its 1999 field shrew study ranged from 4 to 148 mg/kg tPCB ww at sites with a similar range of soil concentrations – i.e., 1 to 42 mg/kg soil (QEA and BBL 2003).

The ERA questions the spatially weighted average soil tPCB concentrations presented in this report and presents alternative concentration values based upon an independent analysis (Vol. 2, p. 10-54; Vol. 6, p. J-54, Table J.4-5). A comparison of the data used in each of the analyses (original and EPA) indicates that the differences are likely due to different methods of calculating spatially weighted average soil tPCB concentrations, different treatments of duplicate sample results, minor differences in assumptions regarding which data points are associated with the grid, revisions to the underlying EPA database that occurred after the original analysis was completed,[23] and EPA's exclusion of nine soil samples from Grid 1 that were collected by GE. In any case, the overall classification of grids as low or high with respect to tPCB concentrations did not change based on EPA's reanalysis of exposure estimates.

EPA also conducted supplemental statistical analyses of the data with the recalculated arithmetic average and spatially weighted average soil tPCB concentrations, using a probit model (Vol. 6, p. J-54). The ERA indicates that these analyses showed a significant, but not strong, relationship between soil tPCB concentrations (both arithmetic and spatially weighted average) and survival of short-tailed shrews from summer to autumn for males, females, and males and females combined (Vol. 2, p. 10-54; Vol. 6, p. J-57). The results of EPA's probit analyses are presented (for both the arithmetic and spatially weighted average tPCB data) for females and males combined (Vol. 6, Figures J.4-11, J.4-12). Other details relating to the results of these analyses are not provided in the ERA.

Dr. Boonstra has reanalyzed the data using the soil tPCB concentrations estimated by EPA and the same statistical method used by EPA (probit analysis). This reanalysis is described in Attachment R to these Comments and summarized in Table 11-1 below. The results of this reanalysis indicate that there was no significant relationship between arithmetic or spatially weighted average concentrations of tPCBs in soil (as calculated by EPA) and survival of shrews (males, females, or males and females combined) from summer to autumn. It is not possible to determine why these results differ from EPA's, because the ERA did not provide all of the data needed to verify the statistical analyses. In any event, Dr. Boonstra's reanalysis of the data confirms that this study provides no evidence of harm to shrews in the PSA, rather than undetermined evidence of harm, as concluded in the ERA (Vol. 2, p. 10-58).[24]

---

[23] Soil tPCB concentrations used in this study were estimated from the March 2002 EPA database (Weston 2002) and the soil concentration maps available online (http://www.epa.gov/region01/ge/thesite/restofriver-maps.html).

[24] The ERA also states that the results of this study were confounded by various factors, such as flooding, habitat differences, potential immigration, and use of body weight to imply fitness (Vol. 2, p. 10-54; Vol. 6, pp. J-58 - J-60). Dr. Boonstra has provided responses to these criticisms in Attachment R.

Table 11-1. Relationship Between the Survival[a] of *Blarina brevicauda* on Six Live-Trapping Grids from Summer to Fall 2001 and Spatially-Weighted Arithmetic tPCB Concentrations (as Calculated in the ERA)

	r	F	p
Males[b]			
Mean Concentrations	-0.739	3.603	0.15
Spatially-Weighted Arithmetic Mean	-0.721	3.247	0.17
Females			
Mean Concentrations	-0.659	3.052	0.16
Spatially-Weighted Arithmetic Mean	-0.53	1.572	0.28
Combined (males+females)[c]			
Mean Concentrations	-0.681	3.451	0.14
Spatially-Weighted Arithmetic Mean	-0.52	1.481	0.29

Notes:
Source of original survival data: Boonstra and Bowman (2003)
a. Survival for each grid was calculated as the probit value.
b. In males, grid 3 was deleted from the analysis because only one male was captured in summer.
c. All grids were included and each animal contributed equally to the analysis.
r = correlation coefficient
F = F-statistic
p = probability level

11.3 Modeled Exposure and Effects

The ERA presents HQs based upon modeled exposure and effects for both the short-tailed shrew and the red fox (Vol. 2, pp. 10-37 - 10-40; Vol. 6, pp. J-49 - J-51). GE has two principal concerns with these analyses.

First, the modeled exposure analyses for both the short-tailed shrew and the red fox used modeled food intake rates developed from a general algorithm for free-living mammals (Vol. 2, pp. 10-17 - 10-18). As noted in prior sections, use of such general algorithms requires assumed inputs on a number of key variables (e.g., free metabolic rate, food preferences, prey assimilation efficiencies, and gross energies) on which limited data are available and to which the results are highly sensitive. This approach is unnecessary because measured food intake rates are available for short-tailed shrews from laboratory studies (Pearson 1947; Morrison et al. 1957; Barrett and Stuek 1976; Ma and Talmadge 2001) and for the red fox from laboratory and captive specimen studies (Vogtsberger and Barrett 1973; Sargeant 1978). Data from these studies should be used in place of the modeled rates to reduce uncertainty, as recommended by EPA's *Wildlife Exposure Factors Handbook* (EPA 1993, p. 3-1).

Second, the effects estimates for the short-tailed shrew and the red fox are based entirely on a very limited number of studies performed using rats and mice (Vol. 2, pp. 10-37 - 10-39; Vol. 6, J-40 - J-42). Given well-documented differences in species sensitivity (e.g., Eisler 1986)**,** this approach produces a high degree of uncertainty in the results.

GE has not identified more appropriate effects thresholds to be used in the HQs for the short-tailed shrew or the red fox. However, the ERA should more fully acknowledge both the uncertainties and the likely over-conservatism of the HQs. For the short-tailed shrew, the over-conservatism of the HQs is illustrated by the fact that they predict high risks from tPCBs in the PSA when the results of the GE-sponsored field study showed no evidence of harm. For the red fox, the large mammal survey was insufficiently quantitative to allow a conclusion regarding evidence of harm, but it did indicate the presence of red fox in the PSA, and the HQs have substantial uncertainties associated with both the exposure and effects inputs. For these reasons, the HQs for both of these species should be accorded Low weight, rather than the Moderate/High weight currently assigned in the ERA (Vol. 2, p. 10-28).

11.4 Overall Assessment

The ERA's weight-of-evidence evaluation assigns all three measurement endpoints – EPA's field studies, GE's short tailed shrew demography study, and the HQs – the same overall Moderate/High weight (Vol. 2, p. 10-58). This evaluation is unwarranted in several respects. First, EPA's field surveys had too many limitations or were not sufficiently quantitative to warrant such a high weight. Second, the weight assigned to GE's population demography study is unduly low, given that it was a site-specific, species-specific, and stressor-specific study that directly addressed the local population-level impacts of exposure to PCBs. Third, the weight assigned to the HQs is unduly high, given the uncertainty associated with the use of generic exposure assumptions and surrogate toxicity data for effects metrics.

The ERA's risk conclusions should likewise be modified. For short-tailed shrews, the ERA concludes that the EPA field survey and the GE shrew demography study provide undetermined evidence of harm, while the HQs predict high-magnitude risks from tPCBs and low risks from TEQs (Vol. 2, p. 10-58). Thus, it concludes that short-tailed shrews are at intermediate risk in the PSA, but qualifies that conclusion as uncertain due to the lack of definitive findings as to whether effects are occurring in the field (Vol. 2, p. 10-63; Vol. 6, p. J-79). However, as discussed above, the GE shrew population demography study showed no evidence of harm, as confirmed by Dr. Boonstra's reanalysis of the shrew survival data using EPA's exposure estimates. Moreover, the HQ analyses are highly uncertain due to the use of a generic food intake rate and rodent effects data. In these circumstances, the weight of the evidence indicates that short-tailed shrews are at low to negligible risk from PCBs in the PSA.

For red fox, the ERA concludes in some places that there are intermediate (although uncertain) risks in the PSA from PCBs and TEQs (Vol. 2, p. 10-63; Vol. 6, p. J-79), and in other places that the evidence of harm and risks for the red fox are undetermined (Vol. 2, pp. 10-59, 10-67; Vol. 6, Table J.4-9). GE concurs with the latter conclusion. The field survey does not provide evidence of harm to the red fox in the PSA; and the HQs are highly uncertain, both because of inappropriate assumptions in the exposure assessment and because of the use of rat data to determine effects metrics. Thus, there are insufficient data to draw defensible conclusions concerning the presence and magnitude of harm or risks to the red fox.

SUMMARY OF CONCLUSIONS
RISKS TO OMNIVOROUS AND CARNIVOROUS MAMMALS

- **The robust field study showing no effects on short-tailed shrews, coupled with the uncertainty and over-conservatism in the model-based HQs, support a conclusion of low or negligible risks to shrews in the PSA.**

- **For red fox, the risks are undetermined, because the only evidence of harm (the HQs) is highly uncertain, and there are thus insufficient data to draw defensible conclusions about risks to red fox in the PSA.**

See Table 11-2 for specific comments on Charge Questions 3.6(a)-(j)

Table 11-2. Assessment Endpoint: Survival, Growth, and Reproduction of Omnivorous and Carnivorous Mammals

EPA Charge Question #3.6	GE Response
(a) Were the EPA studies and analyses performed (e.g., field study, site-specific toxicity study, comparison of exposure and effects) appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), and based on accepted scientific practices?	**EPA Field Surveys:** Basic methodology appears appropriate, but these surveys have limitations. (Section 11.1) **HQs for Short-Tailed Shrew and Red Fox:** Basic methodology is appropriate, but lack of site-specific or species-specific data is a major limitation. (Section 11.3)
(b) Were the GE studies and analyses performed outside of the framework of the ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), based on accepted scientific practices, and incorporated appropriately in the ERA?	**GE Short-Tailed Shrew Demography Study:** This study was appropriate for evaluating whether PCBs are adversely affecting the population demography of short-tailed shrews living in the PSA. The ERA presents a reanalysis of the data from this study. However, a further reanalysis of the data by the principal investigator, using the same methods used in the ERA, continues to show that PCBs are not adversely affecting the short-tailed shrew population. The ERA thus erroneously concludes that this study showed "undetermined" evidence of harm, when in fact it showed no evidence of harm. (Section 11.2)
(c) Were the estimates of exposure appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?	**EPA Field Surveys:** Exposure was not quantified. **GE Short-Tailed Shrew Demography Study:** Exposure was appropriately quantified by developing average soil PCB concentrations for the various grids. EPA's recalculation of the average soil PCBs concentrations for these grids does not change the conclusions of the study. (Section 11.2) **HQs:** Food intake rates should be based on published species-specific measured rates, rather than allometric calculations. (Section 11.3)
(d) Were the effects metrics that were identified and used appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable)?	**GE Short-Tailed Shrew Demography Study:** Effect metrics used are appropriate. (Section 11.2) **HQs:** Effect metrics appear to be derived appropriately, but they are highly uncertain as they are based on surrogate species (i.e., rats and mice). (Section 11.3)
(e) Were the statistical techniques used clearly described, appropriate (i.e., objective, consistent, and reasonable), and properly applied for the objectives of the analysis?	**GE Short-Tailed Shrew Demography Study:** Statistical techniques used in the original study (Attachment Q) are appropriate. The ERA's statistical reanalysis of the data to show an effect on survival could not be replicated by the investigator's statistical analysis using the same inputs and statistical technique used in the ERA. (Section 11.2 and Attachment R) **HQs:** Mathematical calculations appear appropriate.
(f) Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable)?	**Short-Tailed Shrew:** The ERA overstates the risks to shrews because: (1) it mischaracterizes the results of the site-specific field study as showing undetermined evidence of harm when that study showed no evidence of harm: and (2) it gives too much weight to the HQs, which are highly uncertain due to the use of a generic modeled food intake rate and rodent effects data. (Section 11.4) **Red Fox:** The ERA is not clear in its characterization of risk for the red fox. It concludes in some places that there are intermediate (although uncertain) risks and in other places that the evidence of harm and risks for the red fox are undetermined. Given the lack of species-specific toxicity data and the highly uncertain HQ, the risks to red fox cannot be determined. (Section 11.4).

Table 11-2. Assessment Endpoint: Survival, Growth, and Reproduction of Omnivorous and Carnivorous Mammals

EPA Charge Question #3.6	GE Response
(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements could be made.	**GE Short-Tailed Shrew Demography Study:** Uncertainties are overestimated due to an apparently erroneous reanalysis of the data and several other unwarranted criticisms of the study. (Section 11.2 and Attachment R) **HQs:** Uncertainties are generally addressed, but not adequately taken into account in assigning weight of Moderate/High to the HQs. (Section 11.4)
(h) Was the weight of evidence analysis appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable)? If not, how could it be improved?	All three measurement endpoints receive the same overall Moderate/High weight. This evaluation is unwarranted because: (1) the EPA field surveys had too many limitations or were not sufficiently quantitative to warrant such a high weight; (2) the GE short-tailed shrew population demography study should have received a higher weight, given that it was a site-specific, species-specific, and stressor-specific study that directly addressed the local population-level impacts of exposure to PCBs; and (3) the weight assigned to the HQs is unduly high, given the uncertainty associated with the use of generic exposure assumptions and surrogate toxicity data for effects metrics. (Section 11.4)
(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?	Not applicable.
(j) In the Panel members' opinion, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?	**Short-Tailed Shrew: The ERA overstates the risk to local shrew populations.** The robust GE site-specific shrew study shows no evidence of harm, and the HQs are highly uncertain. Thus, the data support a conclusion of low or negligible risks to short-tailed shrews in the PSA. (Section 11.4) **Red Fox: The risks to red fox cannot be determined,** because the only evidence of harm (the HQs) is highly uncertain, and there are thus insufficient data to draw defensible conclusions about risks to red fox in the PSA. (Section 11.4)

12. THREATENED AND ENDANGERED SPECIES (QUESTION 3.8)

Key Points

- The ERA's assessment of risks to T&E species relies solely on modeled exposures and effects (i.e., HQs) for three species – bald eagles, American bitterns, and small-footed myotis. Based on these HQs, it concludes that there are high risks to bald eagles and American bitterns due to tPCBs in the PSA and undetermined risks to small-footed myotis.

- The HQs for bald eagles are unnecessarily conservative and uncertain because they rely on non-species-specific food intake rates and effects metrics, when bald eagle-specific information is available from the literature on both of these parameters. Use of the latter information would reduce the HQs approximately 20-fold.

- The HQs for American bitterns rely on inappropriate dose-based effects metrics for chickens and kestrels when more supportable effects metrics are available. Use of the latter would reduce the HQs by at least 12-fold.

- The HQs for small-footed myotis have substantial limitations and uncertainties due to use of exposure and effects assumptions based on other species (tree swallows and rats, respectively).

- **The ERA overstates the risks to these receptors in the PSA** because:

  Ø It fails to recognize the uncertainties associated with reliance on a single measurement endpoint (HQs), which it does recognize for non-T&E species (e.g., ospreys, red fox);

  Ø It utilizes unnecessarily over-conservative assumptions in the HQs;

  Ø It gives too high weight to the HQ results; and

  Ø It gives the impression that the HQs for three receptors provide an evaluation of multiple lines of evidence, when in fact only one type of measurement endpoint was evaluated.

- The ERA's extrapolation of risks to bald eagles downstream of the PSA concludes that there are risks to wintering bald eagles at Rising Pond. That conclusion is unwarranted because:

  Ø The ERA's analysis assumes that bald eagles would obtain 100% of their winter diet from Rising Pond, whereas that impoundment makes up only about 1% of a wintering bald eagle's foraging range; and

  Ø The MATC used in this analysis was derived using a toxicity threshold for eagle eggs that fails to take account of a relevant study.

12. THREATENED AND ENDANGERED SPECIES (Question 3.8)

The ERA evaluates potential risks to T&E species based on modeled exposures and effects (i.e., HQs) for bald eagles, American bitterns, and small-footed myotis. It concludes that all three representative species are at some risk in the PSA due to tPCB and TEQ exposure, with high risks to bald eagles and American bitterns from tPCBs and undetermined risk to small-footed myotis (Vol. 2, p. 11-65). For areas downstream of the PSA, the ERA concludes that there are risks to bald eagles wintering at Rising Pond, but low risks to bald eagles breeding downstream of the PSA (Vol. 2, p. 11-68). GE believes that the HQs substantially overstate potential risks for T&E species. Our chief concerns are discussed in the following subsections, while specific comments related to Questions 3.8(a)-(j) are provided in Table 12-2 (at the end of this section).

12.1 Modeled Exposures and Effects for Bald Eagles

The ERA's calculation of modeled exposures and effects for bald eagles uses certain exposure assumptions and effects metrics that are not specific to bald eagles, when species-specific information is available. This contributes to unwarranted uncertainty and conservatism in the HQs.

On the exposure side, the ERA uses a modeled food intake rate derived through an algorithm that is based on birds in general, rather than bald eagles in particular. As previously discussed for the other species to which such a general algorithm is applied, this model requires inputs on several key factors for which there are limited data, but which are shown in the sensitivity analysis to strongly influence the results (Vol. 2, p. 11-63; Vol. 6, Table K.2-7). However, data on measured food intake rates are available from a study of food consumption by free-living bald eagles in Connecticut (Stalmaster and Gessaman 1984, Craig et al. 1988), which support a food intake rate that is approximately 18% lower than that yielded by the allometric equation. Although the ERA dismisses the reports on this study on the ground that some eagles did not feed exclusively at the established feeding stations (Vol. 2, p. 11-15; Vol. 6, pp. K-12, K-13), use of such measured rates is preferable to use of the modeled rate, as recognized by the EPA (1993) *Wildlife Exposure Factors Handbook.*

On the effects side, the ERA applies a dose-based effects metric for tPCBs of 0.7 mg/kg bw/d, based on the minor effects level reported by Fernie et al. (2001) for American kestrels (7 mg/kg bw/d), with application of a safety factor of 10 to convert from a LOAEL to a NOAEL (Vol. 2, p. 11-40; Vol. 6, p. K-48). To begin with, there is no justification for application of that safety factor. As the ERA recognizes elsewhere, the effects seen in the Fernie et al. (2001) study at an estimated dose of 7 mg/kg bw/d were minor and generally statistically insignificant (Vol. 2, pp. 7-69, 8-34). Hence, if the Fernie et al. (2001)

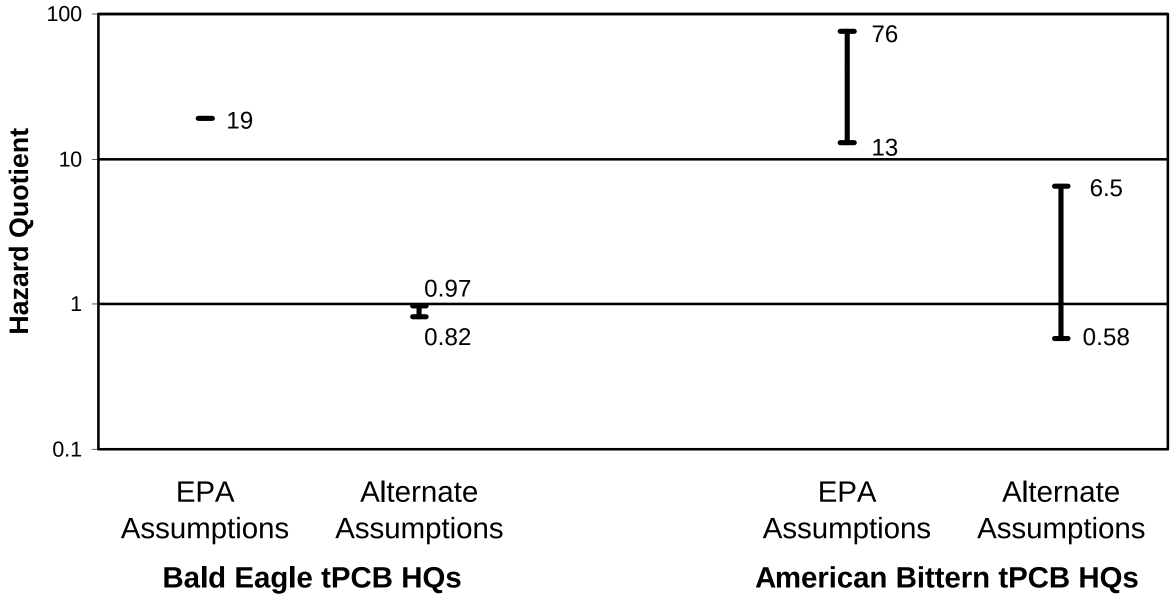
study were applicable, that value should be used directly (i.e., with no safety factor applied), as it is in the effects assessments for American robins, belted kingfishers, and ospreys (Vol. 2, pp. 7-74, 8-35).

Moreover, it would be more accurate and appropriate to derive a dose-based effect metric from studies specific to bald eagles. Stratus (1999) and Donaldson et al. (1999) have developed high quality egg-based effects metrics for tPCBs specific to bald eagles. Indeed, the ERA itself uses the former study to develop a tPCB threshold for bald eagle eggs (Vol. 2, p. 11-41; Vol. 6, p. K-48). The geometric mean value from these two studies is 31.6 mg/kg tPCBs in eggs. Based on that value, a dose-based effects metric of 11 to 13 mg/kg bw/d tPCBs can be back-calculated by applying the same equation employed in the ERA to estimate bald eagle egg concentrations in the PSA, along with Stalmaster and Gessaman's (1984) field-based estimate of food intake rates, and the same assumptions employed in the ERA for maternal transfer, chemical absorption efficiency for tPCBs, and duration of time in the PSA prior to egg-laying. While we recognize that some uncertainty is associated with extrapolations between doses and egg concentrations, the overall certainty in the risk calculations for bald eagles would be substantially increased by the use of a species-specific effects metric.[25]

These recommended changes to the food intake rate and tPCB effects metrics for bald eagles would result in an approximate 20-fold reduction in the HQs for tPCBs. For example, the ERA estimates that the median tPCB dose to bald eagles in the lower PSA is 13 mg/kg bw/d (Vol. 6, Table K.2-6); dividing that value by the ERA's effects metric of 0.7 mg/kg bw/d yields an HQ of 19. However, if the recommended measured food intake rate is used in place of the modeled value, the median dose is reduced to 10.7 mg/kg bw/d; and when that value is divided by the recommended lower-bound effects metric of 11 mg/kg bw/d, the HQ is reduced to 0.97. The impact of these alternative assumptions on the tPCB HQ for bald eagles is illustrated in Figure 12-1 below.

[25] Similarly, in deriving the egg-based effect metric for tPCBs, the ERA should not rely only on the Stratus (1999) study, which suggests a threshold of 20 mg/kg, but should also consider the Donaldson et al. (1999) study, which reported a high quality effect threshold of 50 mg/kg for bald eagle eggs. Thus, the ERA should use the geometric mean of those thresholds, 31.6 mg/kg, as the egg-based effect metric.

Figure 12-1. Impact of Alternative Assumptions on tPCB Hazard Quotients for Threatened and Endangered Birds



Note: Values shown are the upper and lower bound HQs

12.2 Modeled Exposures and Effects for American Bitterns

GE's primary concern related to the development of HQs for American bitterns pertains to the choice of dose-based effects metrics. In the absence of a field-based threshold specific to American bitterns, the ERA employs a tPCB threshold range based on the same studies used for piscivorous birds – 0.12 mg/kg bw/d, based on the Lillie et al. (1974) study of white leghorn chickens (as the most sensitive avian species), to 0.7 mg/kg bw/d, based on the Fernie et al. (2001) study of American kestrels (as a tolerant species) (Vol. 2, p. 11-42; Vol. 6, p. K-66). However, as discussed in our comments on the effects metrics for piscivorous birds (Section 9.3.2), the selection of dose-based effects metrics should consider all relevant studies and should exclude studies conducted on domesticated species. Thus, a more appropriate range of tPCB effects metrics would be 1.4 mg/kg bw/d (from Custer and Heinz [1980] for mallards) to 21.9 mg/kg bw/d (from Custer [2002] for tree swallows), which should be applied in the HQs for American bitterns.[26]

[26] Further, even if the Fernie at al. (2001) study were used for the upper bound of the range, it is not appropriate to apply a 10-fold safety factor to the minor effects level of 7 mg/kg bw/d from that study, for the reasons discussed in Section 12.2 above. Rather, that level should be used directly, as it is for piscivorous birds.

As illustrated in Figure 12-1 above, the recommended changes to the effects metrics for tPCBs applied to American bitterns would result in a reduction in the tPCB HQs for American bitterns by at least 12-fold. For example, for the median tPCB dose to American bitterns in Reach 5A (the site of maximum exposure) (9.07 mg/kg bw/d – Vol. 6, Table K.2-14), use of the alternative recommended effects metric would reduce the tPCB HQ from 76 to 6.5.

For TEQs, as previously discussed for other avian receptors, the selection of effects metrics should not include the Nosek et al. (1992) study based on weekly intraperitoneal injections, but should be based only on the Hoffman et al. (1996) study, which yielded an effects metric of 25,000 ng/kg bw/d. Applying this effects metric to American bitterns would result in a reduction in the HQs for American bitterns by more than three orders of magnitude.

12.3 Modeled Exposures and Effects for Small-Footed Myotis

There are significant limitations in both the exposure assessment and the effects assessment conducted for small-footed myotis. The most important limitation in the exposure assessment is the use of gut content data collected from tree swallows to represent the concentration of COPCs in prey of small-footed myotis. While there is some overlap in the diets of these two species, Table K.2-17 (Vol. 6) of the ERA shows that there are notable differences in the insect classes consumed by the two species, likely reflecting their different foraging strategies. On the effects side, the tPCB effects metric used to evaluate risk to small-footed myotis is based on a dose-response curve developed using effects data from a rat study on TCDD (Vol. 2, p. 11-43; Vol. 6, pp. K-53, K-67), because the available studies on bats (summarized in Vol. 2, pp. 11-36 - 11-39; Vol. 6, pp. K-51 - K-53) focus on bioaccumulation rather than effects, and therefore do not support derivation of effects metrics.

GE has not identified solutions to these limitations in the HQs for small-footed myotis. Rather, these limitations should be more fully acknowledged in the weight-of-evidence evaluation and uncertainty analysis. Given these limitations, the overall weight of Moderate/High given to the HQ results (Vol. 2, pp. 11-59 - 11-62) substantially overstates the quality of these HQs.

12.4 Overall Assessment

Based on the HQ analyses, the ERA concludes that, in the PSA, bald eagles are at high risk due to tPCBs and intermediate risks due to TEQs, that American bitterns are at high risk due to tPCBs and undetermined risks due to TEQs, and that risks to small-footed myotis from tPCBs and TEQs are undetermined (Vol. 2, p. 11-65; Vol. 6, p. K-77). The ERA does not characterize the certainty of these conclusions. While the HQs developed for bald eagles and American bitterns do predict risks, the risk conclusions based on those HQs should be qualified as highly uncertain (if not screening-level), given the availability of only one type of measurement endpoint (i.e., HQs) and the conservatism of that measurement endpoint. For other receptors that were also evaluated solely based on HQs (i.e., ospreys, red fox), the risk summaries acknowledge the uncertainty associated with reliance on a single measurement endpoint.

In addition to failing to recognize the uncertainties inherent in reliance on a single line of evidence, the ERA fails to acknowledge the conservatism of that one line of evidence. For other (non-T&E) receptors that are evaluated based on multiple lines of evidence, the outcomes of higher quality lines of evidence demonstrate the conservatism of the HQs. For example, in the cases of tree swallows, American robins, belted kingfishers, and short-tailed shrews, site-specific field studies show no evidence of adverse reproductive effects despite the prediction of high risks from the HQs. Based on these observations for other receptors, it is reasonable to conclude that the HQs developed for T&E species using the same methodology also likely overstate risks for those receptors. The ERA should recognize this, as it does, for example, for tree swallows and robins (Vol. 2, pp. 7-85, 7-86). In fact, the above alternative analyses indicate that the HQs for bald eagles and American bitterns overestimate risks by at least 20-fold and 12-fold, respectively.

Similarly, the overall weight assigned to the HQs for T&E species (Moderate/High) (Vol. 2, pp. 11-61, 11-62) is overstated in the weight-of-evidence evaluation. Given the HQs' reliance on non-species-specific information in both the exposure assessments and the effects assessments, such weight is much too high. For comparison, the weight-of-evidence evaluations for insectivorous and piscivorous birds consistently apply lower weights to the HQs for those receptors, both in assessing the 10 attributes considered and in the overall weights assigned (Moderate), than do the evaluations for the T&E species, even though the HQs for the latter follow the same approach. This comparison is summarized in Table 12-1 below. While GE believes that even the Moderate weights assigned to the HQs for insectivorous and piscivorous birds are too high, these inconsistencies appear to reflect a bias in the evaluation of the T&E species.

Table 12-1. Comparison of ERA's Weight-of-Evidence Evaluations for Avian HQ for tPCBs

Attributes	Hazard Quotient			
	Tree Swallow and American Robin	Belted Kingfisher and Osprey	Bald Eagle	American Bittern
I. Relationship Between Measurement and Assessment Endpoints				
1. Degree of Association	M	M	H	M/H
2. Stressor/Response	M	M	M/H	M
3. Utility of Measure	M	M	M/H	M/H
II. Data Quality				
4. Data Quality	M	M/H	H	H
III. Study Design				
5. Site Specificity	L/M	L/M	M	M
6. Sensitivity	L/M	L/M	M/H	H
7. Spatial Representativeness	M	M	H	M/H
8. Temporal Representativeness	M	M	M/H	M/H
9. Quantitative Measure	M/H	M/H	H	H
10. Standard Method	M	M	M/H	M/H
Overall Endpoint Value	**M**	**M**	**M/H**	**M/H**
Source of Information	Tables G.4-5, G.4-6	Table H.4-3	Table K.4-3	Table K.4-3

Notes:
L = Low
M = Moderate
H = High

Finally, the ERA's analysis of concurrence among measurement endpoints (Vol. 2, p. 11-62; Vol. 6, pp. K-73 - K-74) gives the impression that the three HQs for three representative receptors provide an evaluation of multiple lines of evidence, when in fact only one type of measurement endpoint was evaluated. If only one type of measurement endpoint is evaluated, the question of concurrence among measurement endpoints becomes a moot point.

SUMMARY OF CONCLUSIONS
<u>RISKS TO T&E SPECIES IN PSA</u>

- **The ERA overstates the risks to bald eagles and American bitterns because it fails to recognize the uncertainties associated with reliance on a single measurement endpoint (HQs) and uses unnecessarily over-conservative assumptions in the HQs. Correction of these assumptions would reduce the HQs by up to a factor of 20. However, any conclusions would still be highly uncertain due to use of only measurement endpoint for each species.**

- **The risks to small-footed myotis, which are also assessed solely on the basis of highly uncertain HQ analyses, are undetermined.**

12.5 Extrapolations of Bald Eagle Risks to Downstream Reaches

The ERA evaluates potential risks to both wintering and breeding bald eagles downstream of the PSA, using different approaches for each. The ERA estimates risks to bald eagles wintering downstream of Woods Pond by comparing concentrations of tPCBs in prey fish collected from Reaches 7 through 16 to a MATC developed for bald eagles. The MATC of 30.41 mg/kg tPCBs in fish (whole body, wet weight) was developed as the dietary concentration (based on the reported winter diet of bald eagles) at which doses of tPCBs to bald eagles would be predicted to exceed the toxicity threshold for eagle eggs (20 mg/kg) (Vol. 2, p. 11-67; Vol. 6, p. K-64). As noted above, the egg-based toxicity threshold was based on a study by Stratus (1999). However, Donaldson et al. (1999) also reported a high quality effect threshold for bald eagle eggs (50 mg/kg), and GE believes that the geometric mean of the thresholds derived from Stratus (1999) (31.6 mg/kg) and Donaldson et al. (1999) should be used to calculate the MATC for bald eagles.

In any event, the ERA indicates that, downstream of the PSA, wintering bald eagles would only be at risk in Reach 8 (Rising Pond) (Vol. 2, p. 11-67; Vol. 6, p. K-64). The ERA describes that finding as conservative, because it assumes that bald eagles would consume fish only from Rising Pond, whereas that pond is considerably smaller than the typical bald eagle foraging area. However, the ERA does not adequately take account of that fact in the risk calculation. In fact, while the wintering foraging range of adult bald eagles is approximately 1,880 ha (Griffin and Baskett 1985), Rising Pond is approximately 18 ha in area (Vol. 1, p. 1-17) – about 1% of the foraging range. It is clearly unrealistic to assume that 100% of a wintering bald eagle's diet could be provided by Rising Pond. Given the very limited availability of suitable foraging area at Rising Pond, either the exposure concentration or the MATC should be adjusted to reflect other likely sources of prey. Correcting these overly conservative assumptions would yield a finding of negligible risks to wintering bald eagles downstream of the PSA.

The ERA uses a different method for assessing risks to breeding bald eagles downstream of the PSA (Vol. 2, p. 11-68; Vol. 6, p. K-65). The ERA estimates risks to bald eagles breeding in Reaches 14 and 15 by estimating doses of tPCBs from fish and by varying assumptions regarding doses received from mammals and waterfowl, for which no tissue data were available. The resultant dose range is then compared to the toxicity threshold of 0.7 mg/kg bw/d developed from the American kestrel study (Fernie et al. 2001). Based on this approach, risks to breeding bald eagles in those two downstream reaches are shown to be below levels of concern. The use of the egg-based MATC, which is applied to wintering bald eagles, would increase the certainty of the conclusion of no risks to breeding bald eagles downstream of Woods Pond.

See Table 12-2 for specific comments on Charge Questions 3.8(a)-(j)

Table 12-2. Assessment Endpoint: Survival, Growth, and Reproduction of Threatened and Endangered Species

EPA Charge Question #3.8	GE Response
(a) Were the EPA studies and analyses performed (e.g., field study, site-specific toxicity study, comparison of exposure and effects) appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), and based on accepted scientific practices?	**HQs for All Receptors (Bald Eagles, American Bitterns, Small-Footed Myotis):** Basic methodology is based on accepted practices. However, several unnecessarily uncertain and overly conservative assumptions are used in the HQs (detailed below).
(b) Were the GE studies and analyses performed outside of the framework of the ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), based on accepted scientific practices, and incorporated appropriately in the ERA?	Not applicable.
(c) Were the estimates of exposure appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable), and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?	**Eagle HQ:** Food intake rates should be based on measured species-specific values reported in *Wildlife Exposure Factors Handbook*, rather than on allometric modeling for birds in general. (Section 12.1) **Myotis HQ:** Uncertainty associated with use of tree swallow gut content data to estimate concentrations of COPCs in myotis prey should be recognized, given the differences in the insect classes consumed by the two species. (Section 12.3)
(d) Were the effects metrics that were identified and used appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable)?	**Eagle HQ:** Dose-based effects metric based on a study of American kestrels is inappropriate because it applies an unjustified 10-fold safety factor and relies unnecessarily on non-species-specific data. Instead, a dose-based effect metric should be derived from bald eagle egg studies. (Section 12.1) **Bittern HQ:** The effects metrics are unsupported. For tPCBs, the upper-bound effects metric should be based on the site-specific tree swallow study, rather than the study of American kestrels, and the lower-bound effects metric should be based on a study on mallards, rather than a study of domesticated chickens. For TEQs, a single effects metric based on Hoffman et al. (1996) should be used, because no other studies are available that offer defensible TEQ effects metrics for birds. (Section 12.2) **Myotis HQ:** Uncertainty associated with use of effects metrics based on rats should be more fully acknowledged in the ERA. (Section 12.3)
(e) Were the statistical techniques used clearly described, appropriate (i.e., objective, consistent, and reasonable), and properly applied for the objectives of the analysis?	**HQs for All Receptors:** Mathematical calculations appear appropriate.
(f) Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable)?	**HQs for All Receptors:** Overly conservative assumptions regarding exposure and effects yield results that overstate risks – e.g., by 20-fold for bald eagles and 12-fold for American bitterns in the tPCB HQs. Changes recommended above would yield more realistic characterization of risk.

Table 12-2. Assessment Endpoint: Survival, Growth, and Reproduction of Threatened and Endangered Species

EPA Charge Question #3.8	GE Response
(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements could be made.	**HQs for All Receptors:** The ERA fails to recognize the uncertainties inherent in reliance on a single line of evidence (HQs) and fails to acknowledge the conservatism of that one line of evidence. In fact, the ERA's analysis of concurrence among measurement endpoints gives the inaccurate impression that HQs for three receptors provide an evaluation of multiple lines of evidence, when in fact only one type of measurement endpoint was evaluated. (Section 12.4).
(h) Was the weight of evidence analysis appropriate under the evaluation criteria (i.e., objective, consistent, and reasonable)? If not, how could it be improved?	**HQs for All Receptors:** Overall weight assigned to HQs (Moderate/High) is overstated. Lack of species-specificity, stressor-specificity, and biological linkage between measurement endpoint and assessment endpoint should be reflected in lower weights. (Section 12.4).
(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?	**Eagles:** The ERA's extrapolation of risks to wintering bald eagles downstream of the PSA overstates risks because: (1) it assumes that bald eagles would obtain 100% of their winter diet from Rising Pond, whereas that impoundment makes up only about 1% of a wintering bald eagle's required foraging range; and (2) the MATC used in this analysis was derived using an eagle egg toxicity threshold that fails to take account of a relevant study. (Section 12.5)
(j) In the Panel members' opinion, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?	**The ERA overstates risks to T&E species** because it fails to recognize the uncertainties associated with reliance on a single measurement endpoint (HQs), and it utilizes unnecessarily over-conservative assumptions in the HQs. Correction of those assumptions would reduce the HQs for bald eagles and American bitterns by up to 20 times, but any conclusions would still be highly uncertain due to use of only one measurement endpoint for each species. For small-footed myotis, the HQs are so uncertain that risks to that species cannot be determined. (Section 12.4).

13. DISCUSSIONS AND CONCLUSIONS (Question 4)

Key Points Relating to Section 12 of the ERA

- The ERA frequently gives too low weight to site-specific field studies that provide direct evidence of whether local populations or communities of ecological receptors in the PSA are being affected by PCBs, and/or it asserts that these field studies provide undetermined evidence of harm when in fact they show no evidence of harm.

- Where field studies were performed by both EPA and GE, the ERA consistently gives greater weight to EPA's studies over GE's studies, regardless of the quality of the study and the data and without adequate justification.

- The ERA overemphasizes the importance of the HQs by making comparisons across receptors based solely on the HQs, without taking into account either other lines of evidence or differences in quality or certainty among the HQs. It also erroneously states that the HQs are not conservative since they generally did not use safety factors. There are numerous other sources of conservatism in the HQs.

- The ERA's extrapolations of potential risks to reaches downstream of the PSA are overstated because they are based on overly conservative MATCs and, in the case of amphibians and bald eagles, fail to take account of habitat limitations downstream of the PSA.

- The ERA's extrapolations of risks from the studied representative species to other species are unsupported by any data on those species, fail to take account of inter-species differences in toxicological sensitivity to PCBs and TEQs, and are based on the highest predicted risk (HQ values), not the overall risk conclusion, for each representative receptor.

- The discussion of various theoretical factors that could result in populations of receptors being abundant despite adverse effects of PCBs appears to have been inserted in an attempt to explain away the findings of abundant populations and diverse communities in the PSA. This section is entirely speculative, unsupported by any studies in the ERA, and has no place in the ERA.

- The general discussion of uncertainties identifies most sources of uncertainty. However, uncertainties associated with field studies are overstated and those associated with HQ analyses are understated.

Overall Conclusions

- **While the ERA includes site-specific studies and multiple lines of evidence for many assessment endpoints, as well as use of a weight-of-evidence approach, it has many substantial flaws, including overemphasis on individual-level (rather than population-level) effects and on HQs, underemphasis of site-specific field studies, incorrect interpretations or weighting of several of the studies, and a general tendency to interpret the lines of evidence toward a finding of risk, even when the data do not support that interpretation.**

- **As a result, the ERA substantially overstates the risks of PCBs to ecological receptors and is not objective, consistent, or reasonable (as specified in the Peer Review Charge) or accurate, reliable, and unbiased (per EPA guidelines). If these flaws are not corrected, GE believes that the ERA cannot serve as a supportable basis for making a remedial action decision for this site.**

13. DISCUSSIONS AND CONCLUSIONS (Question 4)

Section 12 of the ERA (in Volume 2) presents a summary of findings and conclusions from the evaluations of individual assessment endpoints. In addition, Section 12 (a) assesses potential risks of tPCBs to biota in areas downstream of Woods Pond and to species other than the representative receptors considered in Sections 3 through 13, (b) discusses the ecological implications of the assessment results, and (c) summarizes sources of uncertainty that influenced the ERA. As discussed below, the ERA's findings and conclusions frequently overstate the evidence and magnitude of risk and the certainty of conclusions concerning risk. The downstream and interspecies extrapolations are highly conservative and uncertain, and the discussion of "ecological implications" is speculative and unsupported by the data. In general, in making judgments about the interpretation and weighting of the various lines of evidence, the ERA has in most cases adopted the interpretation or weighting that would lead to a conclusion of risk, even when the data do not support that interpretation. This cannot be considered objective, as defined in EPA's (2002a) Information Quality Guidelines as "accurate, reliable, and unbiased."

13.1 Inconsistent Weight-of-Evidence Evaluations

As discussed previously, well-designed, site-specific field studies provide direct evidence concerning whether local populations or communities are being affected by PCBs, and where their results conflict with predictions derived from HQs, such field studies should generally be given greater weight. In a number of cases, however, as described in prior sections of these Comments, the ERA gives too little weight to the field studies or downplays them by mistakenly concluding that they showed "undetermined" evidence of harm. That approach is carried through to the ERA's conclusions, which underweigh site-specific field studies relative to other lines of evidence and/or assert that these studies provide undetermined evidence of harm when in fact they showed no evidence of harm. For example:

- Both the EPA and the GE fish population and community surveys found no evidence of any adverse impacts due to PCBs, yet both studies are said to provide undetermined evidence of harm (p. 12-10), rather than no evidence of harm. Moreover, these two studies are accorded Low/Moderate weights, whereas lines of evidence based on toxicity tests or comparison to literature-derived values are accorded Moderate, Moderate/High, or High weights (p. 12-10).

- For amphibians, all of the literature- and laboratory-derived lines of evidence are accorded moderate or Moderate/High weight, whereas all of the site-specific field studies (except EPA's vernal pond community study) are accorded Low or Low/Moderate weight and are said to have shown undetermined evidence of harm when they found no evidence of harm (p. 12-8).

- EPA's benthic invertebrate community field studies are accorded lower weight than the toxicity tests and are said to provide evidence of harm for coarse sediments (p. 12-7), when a re-evaluation of the data shows no evidence of harm, even for coarse sediments.

- In the case of shrews, both EPA's field study of placental scars and GE's field population demography study, which found no evidence of harm, are said to provide undetermined evidence of harm (p. 12-15). Moreover, the GE study, which directly examined population status over a range of PCB exposures, is accorded no higher weight than literature-derived HQs (p. 12-15).

Only the tree swallow, American robin, and belted kingfisher site-specific reproduction studies are correctly characterized as showing no evidence of harm and are weighted higher than the corresponding literature-based lines of evidence (pp. 12-10, 12-12).

In addition, as discussed throughout these comments, where both EPA and GE contractors performed field studies, EPA's studies are consistently given greater weight, irrespective of the quality of the study and the data. For example:

- In the case of amphibians, EPA's vernal pond community survey is given Moderate/High weight, compared to a Low/Moderate weight for GE's leopard frog egg mass survey, even though both studies are descriptive and neither involved comparisons between exposed and reference populations (p. 12-8). GE's egg mass survey, however, covered 44 ponds within the PSA, whereas EPA's survey covered only 4 ponds. Even EPA's *anecdotal* observations, which do not include any kind of quantitative data, are given a Moderate weight, whereas GE's juvenile wood frog study, which involved statistically rigorous comparisons between larvae with different exposure histories, is accorded a Low weight and completely discounted in the weight-of-evidence evaluation.

- In the case of mink/otter studies, EPA's tracking survey is assigned a higher weight than GE's tracking survey (p. 12-14), even though both studies utilized the same methods and even though GE's survey covered a longer period of time and has better documentation.

- For insectivorous birds, the EPA field study of tree swallow is accorded higher weight than GE's study of American robins (p. 12-10), although the latter is at least of equally high quality.

- In the case of fish and short-tailed shrews, GE's and EPA's studies are assigned equal weights (pp. 12-10, 12-15), even though GE's studies were much more detailed and, in the case of shrews, involved quantitative comparisons among populations exposed to different soil PCB concentrations.

The ERA's ratings of magnitude of harm for the various studies similarly show no apparent logical basis. Why, for example, are the magnitudes of response for GE's short-tailed shrew study intermediate and the EPA's field study low, when both of them are given an undetermined rating with respect to evidence of harm? Why are the average HQ values for fish (between 1 and 10) given ratings of low magnitude of response, while HQ values in the same range for benthic invertebrates are given a rating of high magnitude of harm? No logical explanations are presented for these apparent inconsistencies in the ERA.

Finally, the language used in the ERA to characterize the overall results of the weight-of-evidence analysis is imprecise and, in some cases, inappropriately value-laden. Risks to benthic invertebrates and amphibians are characterized as "significant" (Table 12.2-1), although this term is never defined and no criteria are provided for distinguishing between "significant" and "insignificant" risks. For fish, the ERA provides ambiguous conclusions, stating that a "significant potential risk" is present, that effects on fish are "not severe," and that fish populations in the PSA "are not experiencing catastrophic effects" (Table 12.2-1). It is unclear how a "significant" risk differs from a "significant potential" risk. Moreover, characterizing the risks to be fish as being non-severe and non-catastrophic seems to imply that effects just short of severe or catastrophic might be occurring, even though EPA and GE field studies found no evidence of any effects of PCBs on fish populations in the PSA. In addition, both in the text box on p. 12-1 and in the ERA Conclusions on p. 12-79, PCB-related risks are characterized as "unacceptable." This characterization is inappropriate, because the use of the term "unacceptable" implies a need for management action to reduce risks to "acceptable" levels, and criteria for determining remedial action requirements for the PSA have not been defined.

In short, the complexity and subjectivity of the weight-of-evidence procedure greatly obfuscate any relationships between the results of the studies and the overall risk conclusions. Many of these conclusions are expressed in result-oriented, ill-defined language and do not follow in any obvious way from the weight-of-evidence evaluation. This is not consistent with the requirement in EPA's (2002a) Information Quality Guidelines to present information in an accurate, clear, complete, and unbiased manner.

INCONSISTENCIES IN WEIGHT-OF-EVIDENCE EVALUATION

- **Field studies are underweighed relative to other lines of evidence.**
- **GE field studies are underweighed relative to EPA field studies of the same endpoints.**
- **Ratings of "magnitude of harm" lack a logical basis.**
- **Risks are characterized using imprecise and value-laden terms.**

13.2 Overemphasis on HQs

Although the weight-of-evidence evaluations are intended to be the primary approach for interpreting the various lines of evidence, Section 12.2.2 compares HQs across all receptors. Results of the HQ comparisons are provided graphically in Figures 12.2-1 and 12.2-2. These figures show probabilistic box and whisker plots derived from the probability bounds analysis and Monte Carlo analysis for each endpoint. The plots, and the underlying analyses, convey a misleading impression concerning the absolute and relative risks posed by PCBs and TEQs to the various endpoints considered.

These figures encourage readers to draw conclusions concerning the risks of PCBs present in the PSA from only a single line of evidence, the HQs. All of the field studies are ignored, and many of the HQs are based on literature-derived toxicity data (even when site-specific effects metrics are available). HQs based on species-specific toxicity tests (e.g., wood frog, mink, tree swallow) are compared on the same scale as HQs based on surrogate species such as chickens and rats (robin, kingfisher, fox, short-tailed shrew), giving no indication of the relative certainty of the different HQs. Species for which safety factors were used (eagles, bitterns) are compared on the same scale with other species for which no safety factors were used, with no indication of the differences.

Even more significantly, the ERA's statements that the HQs are "not conservative" and that thus "HQ exceedances of 1 are cause for concern" (pp. 12-21, 12-23) are misleading and, again, inconsistent with the principle of a weight-of-evidence evaluation. According to the ERA, the HQs are not conservative because they did not use safety factors (except for T&E species) and either used thresholds demonstrated to cause effects or used probabilistic techniques to take account of uncertainties (pp. 12-21, 12-23). However, these are not the only possible sources of conservatism in the calculation of HQs. Other sources of conservatism include the interpretation of effects thresholds and the estimation of exposure, as discussed in Section 4.2 and the previous sections on individual receptors. For wildlife species, for example, many of the effects thresholds and exposure assumptions are still highly conservative, even when safety factors are not used.

Moreover, the mere fact that an exposure concentration exceeds an effects threshold – even one based on a species-specific and site-specific study – reveals little or nothing about the severity of the effects that might be expected in the field. It is for this reason that field studies of population- and community-level endpoints are included in assessment approaches such as the Sediment Quality Triad. As noted above, field studies performed for this ERA consistently failed to find effects, even when effects are predicted by HQs. Yet the ERA still indicates that tPCB exposures to species such as robins, belted kingfishers, and short-tailed shrews are of "concern" (pp. 12-23, 12-24), even though the HQs for all of these species are derived from tests performed on surrogate species and conflict with well-designed species-specific field studies that found no evidence of harm.

Further, even when HQs are considered by themselves, HQs above 1 are not necessarily a cause for concern. While HQs below 1 indicate that risks are negligible (rather than "low," as sometimes stated in the ERA), those above 1 indicate a range of potential risks, depending on their magnitude. For example, at the Fort Devens site in Massachusetts, EPA Region 1 has approved reports characterizing HQs between 1 and 10 as indicating "low" potential risks, HQs between 10 and 100 as indicating "moderate" potential risks, and HQs equal to or greater than 100 as indicating "high" potential risks (Arthur D. Little 1995; ARCADIS 2003).

13.3 Extrapolations to Downstream Reaches

The ERA extrapolates potential risks from the PSA to reaches downstream of Woods Pond, for seven of the representative receptors (pp. 12-26 - 12-42). These extrapolations were made by comparing tissue or sediment concentrations of tPCBs to the MATCs derived for those species. GE's concerns with these extrapolations have been discussed in previous sections and may be summarized as follows:

- First, the MATCs applied for benthic invertebrates, amphibians, fish, mink and otter, and bald eagles are lower than levels indicated by the underlying site-specific studies and the scientific literature, as discussed in Sections 5.2.2 (benthic invertebrates), 6.4 (amphibians), 7.3 (fish), 10.6 (mink and otter), and 12.6 (bald eagles). Because the MATCs are biased low, the aerial extent of predicted risks downstream of Woods Pond is overestimated.

- Second, some of the downstream extrapolations are unrealistic due to the ERA's failure to take account of the available habitat for the receptors. For example, although the MATC for amphibians was developed based on sediment tPCB concentrations in breeding ponds, it is applied downstream of Woods Pond to soil tPCB concentrations throughout the entire 100-year floodplain. As detailed in Section 6.6 above, the data on the locations and sizes of vernal pools and wetlands that are available through MassGIS should be used to restrict the amphibian MATC comparisons to only those parts of the floodplain with habitat suitable for amphibian larvae. In the case of bald eagles, the ERA concludes that risks to wintering bald eagles may occur at Reach 8 (Rising Pond), based on the assumption that eagles present in that reach obtain all their food from Rising Pond, even though that impoundment provides only 1% of the surface area necessary to support a single bald eagle (see Section 12.6). Although the ERA recognizes the conservatism in this assumption, no adjustments have been made to the exposure concentrations or the MATC to account for the very small proportion of a bald eagle's diet that could be provided by Rising Pond.

13.4 Extrapolations to Other Species

The ERA also attempts to extrapolate risks from the representative species addressed in the ERA to other species within the same feeding guild, based on factors that influence exposure (e.g., diet, foraging range, body weight) (pp. 12-47 - 12-69). There are a number of problems with this approach:

- First, in the absence of data on the other species, the extrapolations are not supported by any actual evidence.

- Second, although the ERA acknowledges that factors influencing the relative risks of different species include both exposure-related factors and effects-related factors (p. 12-46), the inter-species extrapolations consider only exposure-related factors. They thus do not take account of inter-species differences in sensitivity to the toxicological effects of PCBs and TEQs, which can vary by orders of magnitude even for species within the same feeding guild (e.g., Brunstrom and Lund 1988).

- Third, the comparisons are inherently conservative because, for each representative species, the risk category used for the comparison is assigned based on the highest predicted risk to the receptor, rather than the overall risk conclusion for the receptor. For example, for the purpose of extrapolation to other species, risks to tree swallows, robins, and short-tailed shrews are assumed to be high (based on HQs), even though field studies demonstrated no risks to these species and even though the ERA concludes that tree swallows and robins are not at risk in the PSA. The extrapolation also fails to acknowledge the high uncertainty of the other HQs used in the ERA.

- Fourth, the extrapolations do not account for field data on the other species, where they exist. For instance, in the case of the great blue heron, for which a high risk is assessed based on a comparison to the American bittern, available field data (see Attachment M to these Comments) demonstrate the presence of a highly productive population in the vicinity of the PSA.

13.5 Ecological Implications

Section 12.4.2.2 of the ERA (pp. 12-71 - 12-73) discusses "ecological implications" relating to the interpretation of the risk estimates. This section lists a number of factors that it asserts indicate that studies showing an abundant population of a given receptor at the site do not demonstrate a lack of adverse effects of PCBs on that population. The ERA speculates that: (1) removal of predators (e.g., mink, river otters) could be allowing prey populations to remain abundant, even though these populations may be suffering from effects of PCB exposures, and thus compensates for direct effects of the PCBs; (2) immigration of individuals from uncontaminated areas could be compensating for losses due to PCB exposures; (3) exposure to PCBs may be making some populations more vulnerable to other stressors in the future, even if no adverse effects can be observed at present; (4) PCB exposures may be reducing the genetic diversity of populations, thereby making them less resilient and more susceptible to other stressors; and (5) PCBs might be compromising the immune systems of exposed organisms, making them more vulnerable to infections. The ERA suggests, for example, that removal of predators might be responsible for the failure of field studies to detect effects of PCBs on largemouth bass and short-tailed shrew populations and that immigration is the explanation for sightings of mink tracks in the PSA during winter. Further, the ERA states, in its Conclusions (p. 12-79), that "there are likely indirect effects (e.g., changes in predator-prey relationships, changes in metapopulation dynamics) occurring inside and outside the PSA as a result of the direct impacts caused by tPCBs and other COCs."

The application of these theories to the Rest of River site is entirely speculative and unsupported by any of the underlying studies. While these types of indirect effects of chemical exposures may, in theory, be possible, they certainly are not an inevitable consequence of PCB exposures. An alternative explanation is that the existing levels of exposure are insufficient to cause any of the effects mentioned in the ERA or that density-dependent mechanisms are fully compensating for any effects on individual organisms. It could well be that the abundance of predatory fish, such as largemouth bass, is controlled by the abundance of prey species, and is not measurably affected by the presence or absence of river otters. It could also be that short-tailed shrew and mink populations found in the PSA are self-sustaining, and are not being maintained by immigration. These issues are not addressed in the conceptual model and none of the studies in the ERA was designed to evaluate these types of indirect effects. Indeed, it might be inferred that this entire section is intended to explain away the clearly documented presence of abundant populations and diverse communities in the PSA. In any case, the ERA offers no data to support these theories, and as a result this discussion of these potential indirect effects is pure conjecture. Consequently, these statements are not objective (i.e., accurate, reliable, and unbiased) and should be deleted from the ERA.

13.6 Overall Uncertainty Analysis

The ERA discusses sources of uncertainty affecting each phase of the assessment process (pp. 12-74 - 12-78). In general, while a number of the major sources of uncertainty are identified, the discussions of uncertainties associated with specific receptors or data sets are, in many cases, superficial and inaccurate. For example, the ERA describes the fish community evaluation as "confounded by large habitat variations combined with small overall gradients and large small-scale variation in PCB concentrations," which "made derivations of concentration-response relationships unfeasible" and "limited the studies to qualitative assessments" (p. 12-78). The same characterization could be applied to many population- and community-level field studies, whose purpose is to characterize the overall status of populations and communities, rather than to develop concentration-response relationships. The fact that these studies integrate exposures over large areas and long time scales is a strength, not a weakness, because the populations and communities themselves integrated exposures over the same scales. Population and community studies are not inherently less certain than toxicity tests or individual-level reproduction studies that provide concentration-response relationships – they simply are affected by different uncertainties.

In contrast, the ERA fails to acknowledge the limitations and uncertainties of HQs based on limited numbers of toxicity tests performed on surrogate species. It does not discuss the relative uncertainty of the HQs for different receptor species resulting from differences in quantity and applicability of available toxicity data. In addition, the ERA fails to discuss the high degree of uncertainty associated with extrapolation to downstream reaches and to other species.

SPECULATIVE DISCUSSION OF INDIRECT EFFECTS AND UNBALANCED DISCUSSION OF UNCERTAINTY

- **Speculations concerning indirect effects of PCBs are unsupported by any data and appear to have been intended to explain away results of field studies.**
- **Discussion of uncertainties overstates uncertainties in population and community studies and understates uncertainties in HQ analyses**.

13.7 Overall Conclusions

The ERA includes many site-specific field studies and evaluates most of the assessment endpoints based on multiple lines of evidence. The ERA also uses a systematic weight-of-evidence approach to combine these lines of evidence. Nonetheless, the ERA has some substantial flaws, and as a result of these flaws, it overstates the risks of PCBs to ecological receptors in the Rest of River area. Specifically, the ERA overemphasizes individual-level effects as opposed to effects on local populations and communities, places too much weight on unduly conservative HQs, incorrectly interprets results of some of the toxicity and field studies, and in many cases interprets and applies the results from available field studies in an incorrect and biased manner. The ERA also extrapolates results of the above studies to downstream reaches and other species without acknowledging either the likely overestimates of risk or the very high uncertainty involved in the extrapolations. Finally, the ERA minimizes and attempts to explain away the results of studies performed by both GE and EPA which demonstrate the presence of abundant populations and diverse communities in the PSA. In short, in making judgments about the interpretation of weighting of particular lines of evidence, the ERA generally adopts the interpretation or relative weighting that leads to a conclusion of risk, often stretching or misinterpreting the data to do so, and has downplayed the evidence indicating an absence of harm.

For these reasons, the overall conclusions of the ERA substantially overstate the risks of PCBs to local fish and wildlife populations and communities in the Rest of River area, and cannot be considered objective, consistent, or reasonable (as specified in the Peer Review Charge) or accurate, reliable, or unbiased (as required by EPA's [2002a] Information Quality Guidelines). Indeed, if EPA's risk estimates were correct and applied to the local populations, one would expect, given the length of time that PCBs have been present at elevated levels in the system, that population-level impacts would be evident in the field. Yet the field data show no such impacts.

In these comments, we suggest many changes to the ERA that would address these concerns. GE believes that if these concerns are not adequately addressed, the ERA will not present an accurate picture of risks to ecological receptors in the Rest of River area and cannot serve as a supportable basis for making a remedial action decision for this site.

14. REFERENCES

Alexander. G.R. 1977. Food of vertebrate predators on trout waters in north central lower Michigan. *Michigan Academician* 10:181-195.

Allan, J.D. 1995. *Stream Ecology: Structure and Function of Running Waters*. Kluwer Academic Publishers, Dordrecht, Netherlands. 338 pp.

ARCADIS. 2003. *Final Feasibility Study, AOC 50, Devens Reserve Forces Training Area, Devens, Massachusetts.* Prepared for U.S. Army Forces Command. Prepared by ARCADIS G&M, Inc., Langhorne, Pennsylvania. January.

Arthur D. Little, Inc. 1995. *Final Remedial Investigation Report, AOC-11 (Lovell Road Debris Disposal Area) Fort Devens, Massachusetts*. Prepared for U.S. Army Environmental Center. Prepared by Arthur D. Little, Inc., Cambridge, Massachusetts. August.

Aulerich, R.J., S.J. Bursian, W.J. Breslin, B.A. Olson, and R.K. Ringer. 1985. Toxicological manifestations of 2,4,5,2'4'5'-, 2,3,6,2',3',6'-, and 3,4,5,3',4',5'-phexachlorobiphenyl and Aroclor® 1254 in mink. *Journal of Toxicology and Environmental Health* 15:63-79.

Barrett, G.W., and K.L. Stuek. 1976. Caloric ingestion rate and assimilation efficiency of the short-tailed shrew, *Blarina brevicauda. Ohio Journal of Science* 76:25-26.

Bernstein P., M. Chamberlain, ARCADIS G&M, BBL Sciences, and Branton Environmental Consulting. 2003. *Evaluation of Piscivorous Mammals-Presence/absence, Distribution and Abundance in the Housatonic River Floodplain*. Prepared on behalf of the General Electric Company. June.

Berven. 1990. Factors affecting population fluctuations in larval and adult stages of the wood frog (*Rana sylvatica*). *Ecology* 71:1599-1608.

Bleavins, M.R., R.J. Aulerich, and R.K. Ringer. 1980. Polychlorinated biphenyls (Aroclors 1016 and 1242): Effects on survival and reproduction in mink and ferrets. *Archives of Environmental Contamination and Toxicology* 9:627-635.

Boonstra, R., and L. Bowman. 2003. Demography of short-tailed shrew populations living on polychlorinated biphenyl-contaminated sites. *Environmental Toxicology and Chemistry* 22:1394-1403.

Bosveld, A.T.C., and M. Van den Berg. 1994. Effects of polychlorinated biphenyls, dibenzo-p-dioxins, and dibenzofurans on fish-eating birds. *Environmental Review* 2:147-166.

Brooks, R.P., and W.J. Davis. 1987. Habitat selection of breeding belted kingfishers (*Ceryle alcyon)*. *American Midland Naturalist* 117:63-70.

Brower and Zar. 1977. *Field and Laboratory Methods for General Ecology*. W.M.C. Brown Company Publishers. Dubuque, Iowa. 193 pp.

Brunstrom, B., and L. Lund. 1988. Differences between chick and turkey embryos in sensitivity to 3,4',4,4'-tetrachlorobiphenyl and in concentration/affinity of the hepatic receptor of 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Comparative Biochemistry and Physiology Part C: Comparative Pharmacology and Toxicology* C91:507-512.

Bursian, S.J., R.J. Aulerich, B. Yamini, and D.E. Tillit. 2002. *Dietary Exposure of Mink to Fish from the Housatonic River: Effects on Reproduction and Survival.* Submitted to Roy F. Weston, Inc. July 3.

Bursian, S.J., R.J. Aulerich, B. Yamini, and D.E. Tillit. 2003. *Dietary Exposure of Mink to Fish from the Housatonic River: Effects on Reproduction and Survival.* Final Report. Submitted to Roy F. Weston, Inc. June 10.

Bursian, S.J., and B. Yamini. 2003. *Dietary Exposure of Mink to Fish from the Housatonic River: Inducement of Mandibular and Maxillary Squamous Cell Proliferation.* Submitted to Weston Solutions, Inc. May 1.

Burton, G.A., Jr. 2001. *Assessment of In Situ Stressors and Sediment Toxicity in the Lower Housatonic River*, Revised Draft Report. Institute of Environmental Quality, Wright State University, Dayton, Ohio. November.

Chapman, D.C. 1992. Failure of gas bladder in striped bass: Effect of selenium toxicity. *Archives of Environmental Contamination and Toxicology* 22:296-299.

Colesante, R.T., N.B. Youmans, and B. Ziolkowski. 1986. Intensive culture of walleye fry with live food and formulated diets. *Progressive Fish Culturist* 50:166-169.

Craig, R.J., E.S. Mitchell, and J.E. Mitchell. 1988. Time and energy budgets of bald eagles wintering along the Connecticut River. *Journal of Field Ornithology* 59:22-32.

Custer, C.M. 2002. *Final Report to U.S. Environmental Protection Agency – Exposure and Effects of Chemical Contaminants on Tree Swallows Nesting Along the Housatonic River, Berkshire Co., Massachusetts, 1998-2000.* U.S. Geologic Survey, Biological Resources Division, LaCrosse, Wisconsin. July 8.

Custer, T.W., and G.H. Heinz. 1980. Reproductive success and nest attentiveness of mallard ducks fed Aroclor 1254. E*nvironmental Pollution* 21:313-318.

Donaldson, G.M., J.L. Shutt, and P. Hunter. 1999. Organochlorine contamination in bald eagle eggs and nestlings from the Canadian Great Lakes. *Archives of Environmental Contamination and Toxicology* 36:70-80.

Duke, T.W., J.I. Lowe, and A.J. Wilson, Jr. 1970. A polychlorinated biphenyl (Aroclor 1254) in the water, sediment, and biota of Escambia Bay, Florida. *Bulletin of Environmental Contamination and Toxicology* 5:171.

Eisler, R. 1986. *Polychlorinated Biphenyl Hazards to Fish, Wildlife, and Invertebrates: A Synoptic Review.* Biological Report 85(1.7). U.S. Department of the Interior, U.S. Fish and Wildlife Service, Laurel, Maryland. April.

EPA. 1993. *Wildlife Exposure Factors Handbook*, Volumes I and II. EPA/600/R-93/187a. U.S. Environmental Protection Agency, Office of Research and Development. December.

EPA. 1995. *Great Lakes Water Quality Initiative Technical Support Document for Wildlife Criteria.* EPA-820-B-95-009. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

EPA. 1996a. *Ecological Effects Test Guidelines. OPPTS 850.1010, Aquatic Invertebrate Acute Toxicity Test, Freshwater Daphnids.* EPA 712-C-96-114. U.S. Environmental Protection Agency, Office of Prevention, Pesticides and Toxic Substances. April.

EPA. 1996b. *Ecological Effects Test Guidelines. OPPTS 850.1075, Fish Acute Toxicity Test, Freshwater and Marine.* EPA 712-C-96-118. U.S. Environmental Protection Agency, Office of Prevention, Pesticides and Toxic Substances. April.

EPA. 1999. *Issuance of Final Guidance: Ecological Risk Assessment and Risk Management Principles for Superfund Sites.* OSWER Directive 9285.7-28 P. U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response. Washington, DC. October.

EPA. 2002a. *Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by the Environmental Protection Agency.* EPA/260R-02-008. U.S. Environmental Protection Agency, Office of Environmental Information. Washington, DC. December.

EPA. 2002b. *Calculating Upper Confidence Limits for Exposure Point Concentrations at Hazardous Waste Sites*. OSWER 9285.6-10. U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response.

EVS. 2003. *Assessment of In Situ Stressors and Sediment Toxicity in the Lower Housatonic River.* Adapted from a study by G.A. Burton. GE/Housatonic River Project, Pittsfield, Massachusetts. EVS Environmental Consultants.

FEL. 2002a. *Final Report – Frog Reproduction and Development Study: 2000 Rana pipiens Reproduction and Development Study*. Fort Environmental Laboratories, Inc., Stillwater, Oklahoma.

FEL. 2002b. *Final Report – Frog Reproduction and Development Study: 2000 Rana sylvatica Vernal Pool Study*. Fort Environmental Laboratories, Inc., Stillwater, Oklahoma.

Fernie, K.J., J.E. Smits, G.R. Bortolotti, and D.M. Bird. 2001. Reproductive success of American kestrels exposed to dietary polychlorinated biphenyls. *Environmental Toxicology and Chemistry* 20:776-781.

Gilbert, M., R. LeClair, Jr., and R. Fortin. 1994. Reproduction of the northern leopard frog in floodplain of the Richlieu River, Quebec, Canada. *Journal of Herpetology* 28:465-470.

Goleman W., J. Carr, and T. Anderson. 2002. Environmentally relevant concentrations of ammonium perchlorate inhibit thyroid function and alter sex ratios in developing *Xenopus laevis*. *Environmental Toxicology and Chemistry* 21:590 -597.

Griffin, C.R., T.S. Baskett. 1985. Food availability and winter range sizes of immature and adult bald eagles. *Journal of Wildlife Management* 49:592-594.

Gutleb, A.C., J. Appelman, M. Bronkhorst, J.H.J. van den Berg, and A.J. Murk. 2000. Effects of oral exposure to polychlorinated biphenyls (PCBs) on the development and metamorphosis of two amphibian species (*Xenopus laevis* and *Rana temporaria*). *Science of the Total Environment* 262:147-157.

Hazelton, P.K., R.J. Robel, and A.D. Dayton. 1984. Preferences and influences of paired food items on energy intake of American robins (*Turdus migratorius*) and gray catbirds (*Dumetella carolinensis*). *Journal of Wildlife Management* 48:198-202.

Heaton, S.N., S.J. Bursian, J.P Giesy, D.E. Tillit, J.A. Render, P.D. Jones, D.A. Verbrugge, T.J. Kubiak, and R.J. Aulerich. 1995. Dietary exposure of mink to carp from Saginaw Bay, Michigan. 1. Effects on reproduction and survival, and the potential risks to wild mink populations. *Archives of Environmental Contamination and Toxicology* 28:334-343.

Henning, M., S. Robinson, K. Jenkins. 2002. *Robin Productivity in the Housatonic River Watershed, Berkshire County, Massachusetts.* Prepared for General Electric Company. ARCADIS G&M, Inc., Portland, Maine. April.

Hine, R.L., B.L. Les, and B.F. Hellmich. 1981. *Leopard Frog Populations and Mortality in Wisconsin, 1974-76.* Technical Bulletin No. 122. Department of Natural Resources, Madison, Wisconsin. 39 pp.

Hochstein, J.R., J.A. Render, S.J. Bursian, and R.J. Aulerich. 2001. Chronic toxicity of dietary 2,3,7,8-tetrachlorodibenzo-p-dioxin in adult female min (*Mustela vison*). *Archives of Environmental Contamination and Toxicology* 35:348-353.

Hoffman, D.J., M.J. Melancon, P.N. Klein, C.P. Rice, J.D. Eisemann, R.K. Hines, J.W. Spann, and G.W. Pendleton. 1996. Developmental toxicity of PCB 126 (3,3,4,45-pentachlorobiphenyl) in nestling American kestrels (*Falco sparverius*). *Fundamental and Applied Toxicology* 34:188-200.

Holenweg Peter, A., and H. Reyer. 2002. Species and sex ratio differences in mixed populations of hybridogenetic water frogs: The influence of pond features. *Ecoscience* 9:1-11.

Hornshaw, T.C., R.J. Aulerich, and H.E. Johnson. 1983. Feeding Great Lakes fish to mink: Effects on mink and accumulation and elimination of PCBs by mink. *Journal of Toxicology and Environmental Health* 11:933-946.

Hurlbert, S.H. 1971. The nonconcept of species diversity: A critique and alternative parameters. *Ecology*. 52:577-586.

Kendell, K. 2001. *Northern Leopard Frog Reintroduction: Raven River – Year 2 (2000)*. Alberta Species at Risk Report No. 13. Alberta Sustainable Resource Development, Fish and Wildlife Service, Edmonton, Alberta, Canada. 43 pp.

Kendell, K. 2002. *Survey Protocol for the Northern Leopard Frog.* Alberta Species at Risk Report No. 43. Alberta Sustainable Resource Development, Fish and Wildlife Division, Resource Status and Assessment Branch, Edmonton, Alberta, Canada. 30 pp.

Lillie, R.J., H.C. Cecil, J. Bitman, and G.F. Fries. 1974. Differences in response of caged white leghorn layers to various polychlorinated biphenyls (PCBs) in the diet. *Poultry Science* 53:726-732.

Long, E.R., D.D. MacDonald, S.L. Smith, and F.D. Calder. 1995. Incidence of adverse biological effects within ranges of chemical concentrations in marine and estuarine sediments. *Environmental Management* 19:81-97.

Lowe, J.I., P.R. Parish, J.M. Patrick, and J. Forester. 1972. Effects of polychlorinated biphenyl Aroclor 1254 on American Oyster *Crassostrea virginica*. *Marine Biology* 17:209-214.

Ma, W-C, and S. Talmadge. 2001. Insectivora. In: R.F. Shore and B.A. Rattner, eds. *Ecotoxicology of Wild Mammals.* John Wiley & Sons, Ltd., Chichester, United Kingdom. pp 122-158.

McCarty, J.P., and A.L. Secord. 1999. Reproductive ecology of tree swallows (*Tachycineta bicolor)* with high levels of polychlorinated biphenyl contamination. *Environmental Toxicology and Chemistry* 18:1433-1439.

MDFW. 1979. *Field Investigation Report: Great Blue Heron Rookery Inventory.* Commonwealth of Massachusetts, Division of Fisheries and Wildlife. October 1.

MDFW. 1980. *Field Investigation Report: Great Blue Heron Rookery Inventory, 1980.* Commonwealth of Massachusetts, Division of Fisheries and Wildlife. June 30.

MDFW. 1981. *Field Investigation Report: Great Blue Heron Rookery Inventory. 1981*. Commonwealth of Massachusetts, Division of Fisheries and Wildlife. June 30.

MDFW. 1982. *Field Investigation Report: Great Blue Heron Rookery Inventory, 1982.* Commonwealth of Massachusetts, Division of Fisheries and Wildlife. July 16.

MDFW. 1983. *Field Investigation Report: Great Blue Heron Rookery Inventory, 1983.* Commonwealth of Massachusetts, Division of Fisheries and Wildlife. July 3.

MDFW. 1984. *Field Investigation Report: great Blue Heron Rookery Inventory Results*. Commonwealth of Massachusetts, Division of Fisheries and Wildlife. October 16.

MDFW. 1985. *Field Investigation Report: Great Blue Heron Rookery Inventory Results.* Commonwealth of Massachusetts, Division of Fisheries and Wildlife. December 16.

MDFW. 1986a. *Field Investigation Report: Great Blue Heron Rookery Inventory Results*. Commonwealth of Massachusetts, Division of Fisheries and Wildlife. February 19.

MDFW. 1986b. *Field Investigation Report: Great Blue Heron Rookery Inventory, 1986.* Commonwealth of Massachusetts, Division of Fisheries and Wildlife. November 25.

MDFW. 1987. *Field Investigation Report: Great Blue Heron Rookery Inventory, 1987.* Commonwealth of Massachusetts, Division of Fisheries and Wildlife. October 9.

MDFW. 1989. *Field Investigation Report: Great Blue Heron Rookery Inventory, 1989.* Commonwealth of Massachusetts, Division of Fisheries and Wildlife.

MDFW. 1991. *Memorandum: 1991 Great Blue Heronry Survey*. Commonwealth of Massachusetts, Division of Fisheries and Wildlife. August 19.

MDFW. 1996. *Memorandum: 1996 Great Blue Heronry Survey*. Commonwealth of Massachusetts, Division of Fisheries and Wildlife. November 1.

Menzie, C., M.H. Henning, J. Cura, K. Finkelstein, J. Gentile, J. Maughan, D. Mitchell, S. Petron, B. Potocki, S. Svirsky, and P. Tyler. 1996. Special report of the Massachusetts Weight-of-Evidence Workgroup: A weight-of-evidence approach for evaluating ecological risks. *Human and Ecological Risk Assessment* 2:277-304.

Merrell, D.J. 1977. *Life History of the Leopard Frog in Minnesota.* Occasional Papers 15. Bell Museum of Natural History, University of Minnesota. 23 pp.

Morrison, P.R., M. Pierce, and F.A. Ryser. 1957. Food consumption and body weight in the masked and short-tailed shrews (genus *Blarina*) in Kansas, Iowa, and Missouri. *Annals of the Carnegie Museum* 51:157-180.

Nimmo, D.R., J. Forester, P.T. Heitmuller, and G.H. Cook. 1974. Accumulation of Aroclor 1254 in grass shrimp [I (*Palaemonetes pugio*)] in laboratory and field exposures. *Bulletin of Environmental Contamination and Toxicology* 11:303-308.

Nosek, J.A., S.R. Craven, J.R. Sullivan, S.S. Hurley, and R.E. Peterson. 1992. Toxicity and reproductive effects of 2,3,7,8-tetrachlorodibenzo-p-dioxin in ring-necked pheasant. *Journal of Toxicology and Environmental Health* 35:187-198.

Pearson, O.P. 1947. The rate of metabolism of some small mammals. *Ecology* 29:127-145.

Pielou. 1977. *Mathematical Ecology*. John Wiley & Sons Publishers. New York. 385 pp.

Platanow, N.S., and L.H. Karstad. 1973. Dietary effects of polychlorinated biphenyls on mink. *Canadian Journal of Comparative Medicine* 37:391-400.

Poole, A.F. 1983. Courtship feeding, clutch size, and egg size in ospreys: A preliminary report. In: D.M. Bird, N.R. Seymour, J.M. Gerrard, eds. *Biology and Management of Bald Eagles and Ospreys.* Harpell Press, St. Anne de Bellevue, Quebec. pp. 243-256.

Prose, B.L. 1985. *Habitat suitability index models: Belted kingfishers*. Biological Report 82 (10.87). U.S. Fish and Wildlife Service.

QEA and BBL. 2003. *Housatonic River – Rest of River RCRA Facility Investigation Report.* Prepared for General Electric Company, Pittsfield, Massachusetts. Prepared by QEA, Montvale, New Jersey, and Blasland, Bouck & Lee, Inc., Syracuse, New York. January.

R2 Resource Consultants. 2002. *Evaluation of Largemouth Bass Habitat, Population Structure and Reproduction in the Upper Housatonic River, Massachusetts.* Prepared on behalf of the General Electric Company. Prepared by R2 Resource Consultants. July.

Reeder, A.L, G.L. Foley, D.K. Nichols, L.G. Hansen, B. Wikoff, S. Faeh, J. Eisold, M.G. Wheeler, R. Warner, J.E. Murphy, and V.R. Beasley. 1998. Forms and prevalence of intersexuality and effects of environmental contaminants on sexuality in cricket frogs (*Acris crepitans*). *Environmental Health Perspectives* 106:261-266.

Render, J.A., S.J. Bursian, D.S. Rosenstein, and R.J. Aulerich. 2001. Squamous epithelial proliferation in the jaws of mink fed diets containing 3,3’4,4’5-pentachlorobiphenyl (PCB 126) or 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD). *Journal of Veterinary and Human Toxicology* 43:22-26.

Render, J.A., J.R. Hochstein, and R.J. Aulerich. 2002. Proliferation of periodontal squamous epithelium in mink fed 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD). *Journal of Veterinary and Human Toxicology* 42:85-86.

Saidapur, S.K., N.P. Gramapurohit, and B.A. Shanbhag. 2001. Effect of sex steroids on gonadal differentiation and sex reversal in the frog, *Rana curtipes*. *General and Comparative Endocrinology* 124:115-123.

Sanders, H.O., and J.H. Chandler. 1972. Biological magnification of a polychlorinated biphenyl (Aroclor 1254) from water by aquatic invertebrates. *Bulletin of Environmental Contamination and Toxicology* 7:257-263.

Sargeant, A.B. 1978. Red fox prey demands and implications to prairie duck production. *Journal of Wildlife Management* 36:225-236.

Savage, W.K., F.W. Quimby, and A.P. DeCaprio. 2002. Lethal and sublethal effects of polychlorinated biphenyls (PCBs) on *Rana sylvatica* tadpoles. *Environmental Toxicology and Chemistry* 21:168-174.

Singh, A.K., A. Singh, and M. Engelhardt. 1997. *The Lognormal Distribution in Environmental Applications*. EPA/600/R-97-006. U.S. Environmental Protection Agency. December.

Skorupa, J.P., and R.L. Hothem. 1985. Consumption of commercially-grown grapes by American robins (*Turdis migratorius)*: A field evaluation of laboratory estimates. *Journal of Field Ornithology* 56:369-378.

Stalmaster, M.V., and J.A. Gessaman. 1984. Ecological energetics and foraging behavior of overwintering bald eagles. *Ecological Monographs* 54:407-428.

Stratus. 1999. *Injuries to Avian Resources, Lower Fox River/Green Bay Natural Resource Damage Assessment.* Final Report. Prepared for U.S. Fish and Wildlife Service, U.S. Department of the Interior, and U.S. Department of Justice. Stratus Consulting, Inc. May 7.

Stebbins, R.C., and N.W. Cohen. 1995. *A Natural History of Amphibians*. Princeton University Press, Princeton, New Jersey.

Suter, G.W. II, L.W. Barnthouse, S.M. Bartell, T. Mill, D. Mackay and S. Peterson. 1993. *Ecological Risk Assessment.* Lewis Publishers, Chelsea, Michigan.

Tillitt, D., D. Papoulias, and D. Buckler. 2001. *Fish Reproductive Health Assessment in PCB Contaminated Regions of the Housatonic River, Massachusetts, USA: Investigations of Causal Linkages Between PCBs and Fish Health. Interim Report of Phase I Studies*. Prepared for U.S. Fish and Wildlife Service, Concord, New Hampshire, and U.S. Environmental Protection Agency, Boston, Massachusetts. October 15.

Tillitt, D., D. Papoulias, and D. Buckler. 2002. *Fish Reproductive Health Assessment in PCB Contaminated Regions of the Housatonic River, Massachusetts, USA: Investigations of Causal Linkages Between PCBs and Fish Health. Interim Report of Phase II Studies.* Prepared for U.S. Fish and Wildlife Service, Concord, New Hampshire, and U.S. Environmental Protection Agency, Boston, Massachusetts. November 30.

Tillitt, D., D. Papoulias, and D. Buckler. 2003a. *Fish Reproductive Health Assessment in PCB Contaminated Regions of the Housatonic River, Massachusetts, USA: Investigations of Causal Linkages Between PCBs and Fish Health. Final Report of Phase I Studies*. Prepared for U.S. Fish and Wildlife Service, Concord, New Hampshire, and U.S. Environmental Protection Agency, Boston, Massachusetts.

Tillitt, D., D. Papoulias, and D. Buckler. 2003b. *Fish Reproductive Health Assessment in PCB Contaminated Regions of the Housatonic River, Massachusetts, USA: Investigations of Causal Linkages Between PCBs and Fish Health. Final Report of Phase II Studies.* Prepared for U.S. Fish and Wildlife Service, Concord, New Hampshire, and U.S. Environmental Protection Agency, Boston, Massachusetts.

Tucker, J.W. 1987. Snook and tarpon snook culture and preliminary evaluation for commercial farming. *Progressive Fish Culturist* 49:49-56.

Van den Berg, M., L. Birnbaum, A.T.C. Bosveld, B. Brunstrom, P. Cook, M. Feeley, J.P. Giesy, A. Hanberg, R. Hasegawa, S.W. Kennedy, T. Kubiak, C. Larsen, F.X. Ralaf van Leeuwen, A.K. Jjien Liem, C. Nolt, R.E. Peterson, L. Poellinger, S. Safe, D. Schrenk, D. Tillitt, M. Tyslind, M. Younges, F. Waern, and T. Zacharewski. 1998. Toxic equivalency factors (TEFs) for PCBs, PCDDs, PCDFs for humans and wildlife. *Environmental Health Perspectives* 106:775-792.

Veit, R.R., and W.R. Petersen. 1993. *Birds of Massachusetts*. Natural History of New England Series, Lincoln, Massachusetts: Massachusetts Audubon Society.

Velduizen-Tsoerkan, M.B., D.A. Holwerda, and D.I. Zandee. 1991. Anoxic Survival Time and Metabolic Parameters as Stress Indices in Sea Mussels Exposed to Cadmium or Polychlorinated Biphenyls. *Archives of Environmental Contamination and Toxicology* 20:259-265.

Vogtsberger, L.M., and G.W. Barrett. 1973. Bioenergetics of captive red foxes. *Journal of Wildlife Management* 37:495-500.

Weston. 2002. *U.S. Environmental Protection Agency Housatonic River Watershed Supplemental Investigation Database*. 030102_usepa_hr_dbase1.mdb. Roy F. Weston, Inc. CD-ROM. June.

Winemiller, K.O., and K.A. Rose. 1992. Patterns of life-history diversification in North American fishes: Implications for population regulation. *Canadian Journal of Fisheries and Aquatic Sciences* 49:2196-2218.

Wiley, J.P. 1997. Feathered flights of fancy. *Smithsonian Magazine*. January.

Woodlot. 2003. *Amphibian Reproductive Success Within Vernal Pools Associated with the Housatonic River, Pittsfield to Lenoxdale, Massachusetts.* Prepared for U.S. Environmental Protection Agency, New England Region, Boston, Massachusetts. Woodlot Alternatives, Inc., Topsham, Maine. May.

Wright, A.H. 1920. *Frogs: Their Natural History and Utilization.* Bureau of Fisheries Document No. 888. Washington Government Printing Office, Washington, DC. 44 pp.

Wright, P.J., and D.E. Tillitt. 1999. Embryotoxicity of Great Lakes lake trout extracts to developing rainbow trout. *Aquatic Toxicology* 47:77-92.

Comments of General Electric Company on the Ecological Risk Assessment for the General Electric/ Housatonic River Site, Rest of River (July 2003 Draft)

Attachments

Prepared by:

BBL Sciences
Petaluma, California

ARCADIS G&M, Inc.
Portland, Maine

Branton Environmental Consulting
Vancouver, British Columbia

LWB Environmental Services
Oak Ridge, Tennessee

On Behalf of:

General Electric Company

September 29, 2003

LIST OF ATTACHMENTS

A. Errors and Inconsistencies in the ERA

Prepared by: ARCADIS G&M, Inc., Portland, ME
Branton Environmental Consulting, Vancouver, BC
BBL Sciences, Petaluma, CA

B. Critique of the Method Used to Calculate the 95% Upper Confidence Limit on the Mean of the Sampling Data

Prepared by: AMEC Earth & Environmental, Portland, ME
ARCADIS G&M, Inc., Portland, ME
BBL Sciences, Reston, VA

C. Evaluation of the Benthic Community Study Presented in the ERA

Prepared by: Scott Cooper, Ph.D., University of California, Santa Barbara, CA
Erik Silldorf, Ph.D., University of California, Santa Barbara, CA
BBL Sciences, Petaluma, CA

D. Evaluation of Effects Metrics Developed from Benthic Invertebrate Bioassay Data

Prepared by: BBL Sciences, Petaluma, CA
Branton Environmental Consulting, Vancouver, BC

E. Evaluation of EPA's Northern Leopard Frog Reproduction and Development Study

Prepared by: William J. Resetarits, Ph.D., Old Dominion University, Norfolk, VA
Branton Environmental Consulting, Vancouver, BC
BBL Sciences, Petaluma, CA

F. Northern Leopard Frog (*Rana pipiens*) Egg Mass Survey

Prepared by: ARCADIS G&M, Inc., Portland, ME
William J. Resetarits, Ph.D., Old Dominion University, Norfolk, VA

G. Evaluation of EPA's Wood Frog Vernal Pool Study and EPA's Wood Frog Population Model

Prepared by: William J. Resetarits, Ph.D., Old Dominion University, Norfolk, VA
Branton Environmental Consulting, Vancouver, BC
BBL Sciences, Petaluma, CA
LWB Environmental Services, Oak Ridge, TN

H. Experimental Analysis of the Context-Dependent Effects of Early Life-Stage PCB Exposure on *Rana Sylvatica*

Prepared by: William J. Resetarits, Ph.D., Old Dominion University, Norfolk, VA
(This Attachment is provided only on disk.)

I. Evaluation of USGS Phase I and Phase II Fish Toxicity Studies and Effects Metrics for Fish

Prepared by: BBL Sciences, Petaluma, CA
Branton Environmental Consulting, Vancouver, BC
LWB Environmental Services, Oak Ridge, TN

J. ***In Situ* Reproduction, Abundance and Growth of Young-of-Year and Adult Largemouth Bass in a Population Exposed to Polychlorinated Biphenyls**

Prepared by: Dudley W. Reiser, R2 Resource Consultants, Inc., Redmond, WA
Emily S. Greenberg, R2 Resource Consultants, Inc., Redmond, WA
Thomas E. Helser, National Marine Fisheries Service, Seattle, WA
Margaret Branton, Branton Environmental Consulting, Vancouver, BC
Kenneth Jenkins, BBL Sciences, Petaluma, CA

K. ***Productivity of American Robins Exposed to Polychlorinated Biphenyls, Housatonic River, Massachusetts, USA*** (Manuscript In Press)

Prepared by: Miranda H. Henning, ARCADIS G&M, Inc., Portland, ME
Scott K. Robinson, University of Illinois at Urbana-Champaign, Champaign, IL
Kelly J. McKay, Hampton, IL
Joseph P. Sullivan, Ardea Consulting, Woodland, CA
Heather Bruckert, ARCADIS G&M, Inc., Millersville, MD

L. Productivity and Density of Belted Kingfishers on the Housatonic River

Prepared by: ARCADIS G&M, Inc., Portland, ME
Robert P. Brooks, State College, PA

(This Attachment is provided only on disk.)

M. Productivity Data for Great Blue Herons Breeding in Massachusetts (1980-1999)

Prepared by: ARCADIS G&M, Inc., Portland, ME

(Exhibit to this Attachment is provided only on disk.)

N. Analysis of Kit Survivability Data from EPA's Mink Feeding Study

Prepared by: Branton Environmental Consulting, Vancouver, BC
BBL Sciences, Petaluma, CA

O. Evaluation of Piscivorous Mammals – Presence/Absence, Distribution, and Abundance in the Housatonic River Floodplain

Prepared by: Paul Bernstein, Spencertown, NY
Michael Chamberlain, Ph.D., Louisiana State University, Baton Rouge, LA
ARCADIS G&M, Inc., Albany, NY
BBL Sciences, Petaluma, CA
Branton Environmental Consulting, Vancouver, BC

(Exhibits to this Attachment are provided only on disk.)

P. Comments Relating to Mink and Otter Field Studies

Prepared by: Michael J. Chamberlain, Ph.D., Louisiana State University, Baton Rouge, LA

Q. ***Demography of Short-Tailed Shrew Populations Living on Polychlorinated Biphenyl-Contaminated Sites***

Prepared by: Rudy Boonstra, Ph.D., University of Toronto, Scarborough, ON
Lanna Bowman, University of Toronto, Scarborough, ON

R. Comments Relating to Short-Tailed Shrew Demography Study

Prepared by: Rudy Boonstra, Ph.D., University of Toronto, Scarborough, ON